

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 17:35:29 ; Search time 723.713 Seconds

(without alignments)
1865.405 Million cell updates/sec

Title: US-09-594-065-4

Perfect score: 33

Sequence: 1 cccctgaactcgtctccctccagatgcaggg 33

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBdb1.*

1: gb_da.*

2: gb_htg.*

3: gb_in.*

4: gb_cm.*

5: gb_ov.*

6: gb_pac.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_da.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pac.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vir.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23.6	71.5	180155	10 AC112273	AC112273 Mus muscu
2	22.6	68.5	2712	6 BD165793	BD165793 Vaccine f
3	22.6	68.5	2715	6 BD145143	BD145143 Herpes si
4	22.6	68.5	2818	14 AF021340	AF021340 Human her
5	22.6	68.5	2846	14 HHU12174	HHU12174 Human herpe
6	22.6	68.5	2846	14 HHU12174	HHU12174 Human herpe
7	22.6	68.5	2855	14 HHU12173	HHU12173 Human herpe
8	22.6	68.5	2855	14 HHU12175	HHU12175 Human herpe
9	22.6	68.5	3324	14 HS2GB	HS2GB
10	22.6	68.5	3472	6 BD165792	BD165792 Vaccine f
11	22.6	68.5	3472	6 HS2GB3	HS2GB3
12	22.6	68.5	3715	14 HS2GB2A	HS2GB2A
13	22.6	68.5	154746	14 HSV2HG52	HSV2HG52
14	22.6	67.9	2281	9 AF181264	AF181264 Homo sapi
15	22.4	67.9	2815	6 AR308082	AR308082 Sequence
16	22.4	67.9	3284	9 BC033100	BC033100 Homo sapi
17	22.4	67.9	3583	9 AF214736	AF214736 Homo sapi
18	22.4	67.9	151949	9 CNS01DS7	CNS01DS7 BAC sequ
19	22.4	66.1	113975	8 OSJN00141	OSJN00141
20	21.8	66.1	208920	2 AC140469	AC140469 Mus muscu
21	21.8	66.1	212390	2 AC133576	AC133576 Mus muscu
22	21.8	66.1	212390	2 AC133576	AC133576 Mus muscu
23	21.4	64.8	798	6 103657	103657 Sequence 5
24	21.4	64.8	1478	14 S74390	S74390 GPath/UL27
25	21.4	64.8	1488	14 S65444	S65444 UL27-glycop
26	21.4	64.8	1857	14 HS1GBM	HS1GBM
27	21.4	64.8	2088	6 E03093	E03093 DNA encodin
28	21.4	64.8	2712	6 BD145144	BD145144 Herpes si
29	21.4	64.8	2715	6 BD165795	BD165795 Vaccine f
30	21.4	64.8	2715	14 AF097023	AF097023 Human her
31	21.4	64.8	2715	14 AF311740	AF311740 Human her
32	21.4	64.8	2715	14 HHU49121	HHU49121 Human herpe
33	21.4	64.8	2809	14 AF259899	AF259899 Human her
34	21.4	64.8	2925	6 E03115	E03115 DNA encodin
35	21.4	64.8	3098	6 E00883	E00883 Fragment of
36	21.4	64.8	3461	6 E01195	E01195 DNA sequenc
37	21.4	64.8	3465	6 E03024	E03024 DNA encodin
38	21.4	64.8	3465	6 E03092	E03092 DNA encodin
39	21.4	64.8	3465	6 E03112	E03112 DNA encodin
40	21.4	64.8	3472	6 BD165794	BD165794 Vaccine f
41	21.4	64.8	3472	6 108686	108686 Sequence 2
42	21.4	64.8	3640	6 108361	108361 Sequence 1
43	21.4	64.8	3643	14 HS1GBP	HS1GBP
44	21.4	64.8	3755	6 E03025	E03025 DNA encodin
45	21.4	64.8	3755	6 E03113	E03113 DNA encodin

ALIGNMENTS

RESULT 1

AC112273 180155 bp DNA linear ROD 29-AUG-2002

LOCUS AC112273 Mus musculus chromosome 3 clone RP24-122H5, complete sequence.

DEFINITION AC112273

ACCESSION AC112273

VERSION AC112273.4 GI:22539282

KEYWORDS HTG.

ORGANISM Mus musculus (house mouse)

REFERENCE 1 (bases 1 to 180155)

AUTHORS Matheron,J.D., and Waterson,R.H.

TITLE The sequence of Mus musculus clone

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 180155)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

REFERENCE 3 (bases 1 to 180155)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 180155)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

COMMENT On Aug 29, 2002 this sequence version replaced gi:21240758.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submission@watson.wustl.edu
----- Project Information -----
Center project name: W_B0122405

FEATURES
Source
1. 180155
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/clone="RP24-122H5"

BASE COUNT 51879 a 36032 c 37337 g 54907 t

ORIGIN

Query Match 71.5%; Score 23.6; DB 10; Length 180155;
Best Local Similarity 81.2%; Pred. No. 29;
Matches 26; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 CCTGCAAACTGCTGCTCTCCAGCATGCAGG 32
DB 103608 CCCAGCAACTGCTGCTGTCACAAACATGCAGG 103639

RESULT 2
LOCUS BD165793 2712 bp DNA linear PAT 17-JAN-2003
DEFINITION Vaccine for use in the therapeutic treatment of HSV.
ACCESSION BD165793
VERSION BD165793.1 GI:27871605
KEYWORDS JP 2002167398-A/2.
SOURCE Herpes simplex virus 7
ORGANISM Herpes simplex virus 7
Virus;; dDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
1 (bases 1 to 2712)
Burke,R.L., Pechl,C. and Valenzuela,P.D.T.
Vaccine for use in the therapeutic treatment of HSV
Patent: JP 2002167398-A 2 11-JUN-2002;
CHIRON CORP

COMMENT OS Herpes simplex virus 7
PN JP 2002167398-A/2
PD 11-JUN-2002
PF 12-SEP-2001 JP 2001277340
PR 20-OCT-1986 US 921.213
PI RAE LYN BURKE,CAROL PACHL,PABLO D T VALENZUELA PC
C07K14/035,A61K39/245,A61P31/22,C12N15/09,C12N15/00,PC
C12N15/00
CC Vaccine for use in the therapeutic treatment of HSV FH Key
FT CDS Location/Qualifiers
1. 2712

FEATURES
Source
1. 2712

/organism="Herpes simplex virus 7"
/mol_type="genomic DNA"
/db_xref="taxon:57277"

BASE COUNT 490 a 938 c 884 g 400 t

ORIGIN

Query Match 68.5%; Score 22.6; DB 6; Length 2712;
Best Local Similarity 95.7%; Pred. No. 90;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 CAAACTGCTGCTCTCCAGCATG 28
DB 2041 CAAACTGCTGCTCTCCAGCATG 2019

RESULT 3
LOCUS BD145143 2715 bp DNA linear PAT 17-JAN-2003
DEFINITION Herpes simplex virus VP16 vaccine.
ACCESSION BD145143
VERSION BD145143.1 GI:27850901
KEYWORDS JP 2002136297-A/3.
SOURCE unidentified
ORGANISM unidentified
unclassified.
1 (bases 1 to 2715)
Burke,R.L. and Sekulovich,R.E.
Herpes simplex virus VP16 vaccine
Patent: JP 2002136297-A 3 14-MAY-2002;
CHIRON CORP

COMMENT OS HSV GB2
PN JP 2002136297-A/3
PD 14-MAY-2002
PF 10-SEP-2001 JP 2001274335
PR 02-AUG-1990 US 561.528
PI RAE LYN BURKE,ROSE B SEKULOVICH
PC C12N15/09,A61K39/00,A61K39/245,A61P31/12,C07K14/03,C12N15/00,PC
A61K37/02
CC Herpes simplex virus VP16 vaccine
FH Key
FT CDS Location/Qualifiers
1. 2715
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 491 a 938 c 886 g 400 t

ORIGIN

Query Match 68.5%; Score 22.6; DB 6; Length 2715;
Best Local Similarity 95.7%; Pred. No. 90;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 CAAACTGCTGCTCTCCAGCATG 28
DB 2041 CAAACTGCTGCTCTCCAGCATG 2019

RESULT 4
LOCUS AF021340 2818 bp DNA linear VRL 24-AUG-2000
DEFINITION Human herpesvirus 2 glycoprotein B precursor (UL27) gene, complete
cde.
ACCESSION AF021340
VERSION AF021340.1 GI:2465440
KEYWORDS Human herpesvirus 2
SOURCE Human herpesvirus 2
ORGANISM Human herpesvirus 2
Virus;; dDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
1 (bases 1 to 2818)
Terhune,S.S., Coleman,K.T., Sekulovich,R., Burke,R.L. and
Spear,P.G.

TITLE Limited variability of glycoprotein gene sequences and neutralizing targets in herpes simplex virus type 2 isolates and stability on passage in cell culture

JOURNAL J. Infect. Dis. 178 (1), 8-15 (1998)

MEDLINE 98314525

PUBMED 9652417

REFERENCE 2 (bases 1 to 2818)

AUTHORS Terhune,S.S., Coleman,K.T., Sekulovich,R., Burke,R.L. and Spear,P.G.

TITLE Direct Submission

JOURNAL Submitted (29-AUG-1997) Microbiology-Immunology, Northwestern University, 303 East Chicago Ave, Chicago, IL 60611, USA

FEATURES Location/Qualifiers

source 1..2818

organism="Human herpesvirus 2"

/mol_type="genomic DNA"

/strain="BHKC"

/db_xref="taxon:10310"

gene 1..2715

/gene="UL27"

/note="UL27"

/codon_start=1

/product="glycoprotein B precursor"

/protein_id="AAB72100.1"

/db_xref="GI:2465441"

/translation="MRGGGLICALVGVAAVAASAPAPAPRAGGVAATVAANGPAPRPVPSPTATTPARKKTKKPPRPATPPDANATVAAGHATLRHLEIKYEMADQFVCPBPATGATVVOEPORCPTREPGNTGIAVFKENIAPRKATRTMYKDTVSGVGHRTYSQPMGI FEDRAPVPEEVIDKINAKGCKSTAKYVNMETTAFFHDEHEDMEIKPAKVAATRSRGMHTTDLKNSRVAEFKRGYTTNCIVEEDARSVPYDEFLATGDEVMSPFYREGSHTEHTSYAARFKVDGFYARDLTAKATSPNLTTPKFTVAMWVPKRPACVMTKQEVDEMRAEYGSFRSSDA1STFTNLTNLYSLSRVLDGICIGDRAEADIMEAKKVAATHIKVGQPOYLAATGGPLIAYOPLLSMTALVYREYREBDKRNATTPAPLRAPASANSVERIKTSSIEPARLOFTNHIORVNDMLGRIVANMCELDQNEHLTMNEARKLNPAIASATVGRVSAKMDVAVSTCVPAEDNVIVQNSMRVSSRPCTYSRPLVSFRYEDQPLIEGQGENNELRRLVDALEPTGVGHRRYFIFGGGVYFEERYASHQLSRADVTYSTFDLNIITMLDEHEFVPLEYTRHEIKDSGLDYTEVQRNQLHDLRFADIDIVIRADANAMAPAGCAFEFGDGLDGAIVKGVGVAVSVGVSFMSNPFGLAVGLVLAGLVAAFPAPFRVVLQORNMKALVLTTEKLTSDPGVSGBGBGAGCGGDEKLABAREMIRYALVSMERHEKARKGTGTSALLSKVTNVLNRKNKARYSPLNDEADGDEDEL"

BASE COUNT 515 a 952 c 920 g 431 t

ORIGIN

Query Match 68.5%; Score 22.6; DB 14; Length 2818;

Best Local Similarity 95.7%; Pred. No. 90;

Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CAAACTGCTGCTCTCCAGCATG 28

Db 2041 CAAACTGCTGCTCTCCAGCATG 2019

RESULT 5 HHU12172/c 2846 bp DNA linear VR1_05-AUG-1994

LOCUS Human herpesvirus 2 CAM4B glycoprotein B (UL27) gene, complete cds.

DEFINITION U12172

ACCESSION U12172

VERSION U12172.1 GI:517447

KEYWORDS Human herpesvirus 2

SOURCE Human herpesvirus 2

ORGANISM Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Simplexvirus.

REFERENCE 1 (bases 1 to 2846)

AUTHORS Terhune,S.S. and Spear,P.G.

TITLE Variability of herpes simplex virus glycoproteins in clinical specimens

JOURNAL Unpublished

REFERENCE 2 (bases 274 to 2846)

AUTHORS Stuve,L.L., Brown-Shimer,S., Pachl,C., Najarian,R., Dina,D. and Burke,R.L.

TITLE Structure and expression of the herpes simplex virus type 2 glycoprotein gB gene

JOURNAL J. Virol. 61 (2), 326-335 (1987)

MEDLINE 87112925

PUBMED 3027364

REFERENCE 3 (bases 1 to 2846)

AUTHORS Spear,P.G.

TITLE Direct Submission

JOURNAL Submitted (11-JUL-1994) Patricia G. Spear, Microbiology-Immunology, Northwestern University Medical School, 303 East Chicago Avenue, Chicago, IL 60611, USA

FEATURES Location/Qualifiers

source 1..2846

organism="Human herpesvirus 2"

/mol_type="genomic DNA"

/strain="CAM4B"

/db_xref="taxon:10310"

gene 36..2741

/gene="UL27"

/note="UL27"

/codon_start=1

/product="glycoprotein B"

/protein_id="AAB60545.1"

/db_xref="GI:522168"

/translation="MRGGGLICALVGVAAVAASAPAPRAGGVAATVAANGPAPRPVPSPTATTPARKKTKKPPRPATPPDANATVAAGHATLRHLEIKYEMADQFVCPBPATGATVVOEPORCPTREPGNTGIAVFKENIAPRKATRTMYKDTVSGVGHRTYSQPMGI FEDRAPVPEEVIDKINAKGCKSTAKYVNMETTAFFHDEHEDMEIKPAKVAATRSRGMHTTDLKNSRVAEFKRGYTTNCIVEEDARSVPYDEFLATGDEVMSPFYREGSHTEHTSYAARFKVDGFYARDLTAKATSPNLTTPKFTVAMWVPKRPACVMTKQEVDEMRAEYGSFRSSDA1STFTNLTNLYSLSRVLDGICIGDRAEADIMEAKKVAATHIKVGQPOYLAATGGPLIAYOPLLSMTALVYREYREBDKRNATTPAPLRAPASANSVERIKTSSIEPARLOFTNHIORVNDMLGRIVANMCELDQNEHLTMNEARKLNPAIASATVGRVSAKMDVAVSTCVPAEDNVIVQNSMRVSSRPCTYSRPLVSFRYEDQPLIEGQGENNELRRLVDALEPTGVGHRRYFIFGGGVYFEERYASHQLSRADVTYSTFDLNIITMLDEHEFVPLEYTRHEIKDSGLDYTEVQRNQLHDLRFADIDIVIRADANAMAPAGCAFEFGDGLDGAIVKGVGVAVSVGVSFMSNPFGLAVGLVLAGLVAAFPAPFRVVLQORNMKALVLTTEKLTSDPGVSGBGBGAGCGGDEKLABAREMIRYALVSMERHEKARKGTGTSALLSKVTNVLNRKNKARYSPLNDEADGDEDEL"

BASE COUNT 522 a 962 c 930 g 432 t

ORIGIN

Query Match 68.5%; Score 22.6; DB 14; Length 2846;

Best Local Similarity 95.7%; Pred. No. 90;

variation 97

/gene="UL27"

/citation=(2)

/replace="cgsgcgcccc"

variation 172

/gene="UL27"

/citation=(2)

/replace="g"

variation 205

/gene="UL27"

/citation=(2)

/replace="a"

variation 876

/gene="UL27"

/citation=(2)

/replace="c"

variation 1015

/gene="UL27"

/citation=(2)

/replace="g"

variation 1250

/gene="UL27"

/citation=(2)

/replace="c"

Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 6 CAAACTCGTGTCTCCAGCATG 28
 |||||:|||||
 Db 2067 CAAACTCGTGTCTCCAGCATG 2045

RESULT 6
 HHU12174/c 2846 bp DNA linear VRL 05-AUG-1994
 LOCUS Human herpesvirus 2 MMA glycoprotein B (UL27) gene, complete cds.
 DEFINITION U12174
 ACCSSION U12174.1 GI:517451
 VERSION
 KEYWORDS
 SOURCE Human herpesvirus 2
 ORGANISM Human herpesvirus 2
 Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 Alphaherpesvirinae; Simplexvirus.
 REFERENCE 1 (bases 1 to 2846)
 Terhune, S.S. and Spear, P.G.
 TITLE Variability of herpes simplex virus glycoproteins in clinical specimens
 JOURNAL Unpublished
 REFERENCE 2 (bases 274 to 2846)
 Stuve, L.L., Brown-Shimer, S., Pachl, C., Najarian, R., Dina, D. and Burke, R.L.
 TITLE Structure and expression of the herpes simplex virus type 2 glycoprotein gB gene
 JOURNAL J. Virol. 61 (2), 326-335 (1987)
 MEDLINE 87112925
 PUBMED 3027364
 REFERENCE 3 (bases 1 to 2846)
 Spear, P.G.
 TITLE Direct Submision
 JOURNAL Submitted (11-JUL-1994) Patricia G. Spear, Microbiology-Immunology, Northwestern University Medical School, 303 East Chicago Avenue, Chicago, IL 60611, USA
 FEATURES
 source
 1. 2846
 /organism="Human herpesvirus 2"
 /mol_type="genomic DNA"
 /strain="MMA"
 /db_xref="taxon:10310"
 36..2741
 /gene="UL27"
 36..2741
 /gene="UL27"
 /note="gb; precursor peptide"
 /codon_start=1
 /product="glycoprotein B"
 /protein_id="AAB60547.1"
 /db_xref="GI:522170"
 /translation="MRGGGLICALVGVLAANAASAPAPRAGSGVAATYAANGRA
 SRPPVSPATIRAKRKTKPKPEPEKATPPDAATVAAGHATLRAHLETKENAD
 AOFVYCPPTATVQFEPQPRCPTRPGQVYTBELAVFENLAIFYKMTAKYKDV
 TVSQVPGHRSQFWGIFEDRAPVFEVEIVIKNAKGCRSTAKVRNMMEITAFHD
 DHEITMELKPAKVAATRTSGMHTDLKTPNPRVEAFHNGTVCIVIEVARSVYR
 DEPLVATGDFVMSPPVYGRSGSHTEHSVAADRFKQVDFVADLTAKATSPTR
 NLITTPKTVAMDWPRKRPAYCTMKQOEDEM.PAYGSGFRSSDAISTPTNLT
 EYSISRVDLCTIGDAREATDRPKRTNMTIKVGPOTYLATGFLIYOPILSN
 TLAEIYREYREDDRKRNATPALEAPASMAVERIKTTSIEFARLQFTVNHQ
 RHVNDMLGRIVAWCELONHLELTLMNEARKLPNAIASATVGRVSAMLGDVAVST
 CVPVAPDNVIVONSMBRVSSRPGTCYSLVFRFYDQPLIEGQLENNELRLDAL
 EPCYVGHRYFLPGGYYFEEYASHLSHADVTTSFTFLDNTIMLDEHPPLVEY
 YTRHFKSGSLDTEVQRNQHRLRADIDTYIRANAMNAPGLCAFEFGCDLG
 RAQKRVNGVGVSAVSVSSFNPFGLAVGLLVLAGLVAAFFAFRYVLQORN
 EMKALYPLTKELKTSIDPGVGVGEGEAGGGEAGGLAARREMIYMALVSAMERT
 EHKARKGTSALLSSKVTNNVLRKKNKARYSPLHNEDAGDEDEL"
 36..101
 /gene="UL27"
 97
 /gene="UL27"
 /citation=[2]
 sig_peptide
 variation

variation
 143
 /replace="cggcggcccc"
 /gene="UL27"
 /citation=[2]
 /replace="c"
 variation
 205
 /gene="UL27"
 /citation=[2]
 /replace="a"
 variation
 237
 /gene="UL27"
 /citation=[2]
 /replace="a"
 variation
 876
 /gene="UL27"
 /citation=[2]
 /replace="t"
 variation
 1015
 /gene="UL27"
 /citation=[2]
 /replace="g"
 variation
 1212
 /gene="UL27"
 /citation=[2]
 /replace="t"
 variation
 1250
 /gene="UL27"
 /citation=[2]
 /replace="c"
 variation
 2642
 /gene="UL27"
 /citation=[2]
 /replace="t"
 variation
 2642
 /gene="UL27"
 /citation=[2]
 /replace="c"
 variation
 959 c 934 g 433 c

BASE COUNT
 520 a 959 c 934 g 433 c
 ORIGIN

Query Match 68.5%; Score 22.6; DB 14; Length 2846;
 Best Local Similarity 95.7%; Pred. No. 90;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 6 CAAACTCGTGTCTCCAGCATG 28
 |||||:|||||
 Db 2067 CAAACTCGTGTCTCCAGCATG 2045

RESULT 7
 HHU12173/c 2855 bp DNA linear VRL 05-AUG-1994
 LOCUS Human herpesvirus 2 JD23 glycoprotein B (UL27) gene, complete cds.
 DEFINITION U12173
 ACCSSION U12173.1 GI:517449
 VERSION
 KEYWORDS
 SOURCE Human herpesvirus 2
 ORGANISM Human herpesvirus 2
 Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 Alphaherpesvirinae; Simplexvirus.
 REFERENCE 1 (bases 1 to 2855)
 Terhune, S.S. and Spear, P.G.
 TITLE Variability of herpes simplex virus glycoproteins in clinical specimens
 JOURNAL Unpublished
 REFERENCE 2 (bases 274 to 2855)
 Stuve, L.L., Brown-Shimer, S., Pachl, C., Najarian, R., Dina, D. and Burke, R.L.
 TITLE Structure and expression of the herpes simplex virus type 2 glycoprotein gB gene
 JOURNAL J. Virol. 61 (2), 326-335 (1987)
 MEDLINE 87112925
 PUBMED 3027364
 REFERENCE 3 (bases 1 to 2855)
 Spear, P.G.
 TITLE Direct Submision
 JOURNAL Submitted (11-JUL-1994) Patricia G. Spear, Microbiology-Immunology,

Northwestern University Medical School, 303 East Chicago Avenue,
Chicago, IL 60611, USA

FEATURES

SOURCE

1. .2855
Location/Qualifiers

/organism="Human herpesvirus 2"

/mol_type="genomic DNA"

/strain="JD23"

/db_xref="taxon:10310"

36. .2750

/gene="UL27"

36. .2750

/note="gb: precursor peptide"

/product="glycoprotein B"

/protein_id="AA860546.1"

/db_xref="GI:522169"

/translation="MRGGSLICLVGALVAANAAPAPAPASGVAATYANG
GPASRPPVPSPATYKARKTKKPPKPPATPPDANATVAAGHATLRALREIKYE
NADAQFYVCPPTGATVVOEPORPCPTREGQYTGIAVFPENIAPYFKATMY
KDVTSQVMFGHRSQFMGI FEDRAVPFEEVIDIKINAGVRSATAYVRNMETAF
HRDDETDLKPAKVATRTSGMHTDLKNSRVAPFRYGTTCVIEVDARSV
YPRDEPLATGDPVYNSPYGREGSTHTSTYAADFKQVDFYARDLTAKATISB
TTRNLTLTPKFTVAMDVPPKPAVCTMTKQVDEMLRAEYGSFRSSDAISTFTT
NLTEYSLSRVDLGDICGRDAREALDRFARKYNATHIKVGPQYTLATGFLIAYOPL
LSNTLAEIYVREYREBODRKPRNATPAPLEAPASANSVERIKTSSIEFARLOFTYN
HIORHVDMLGRIAVANCELOHNETLTMEARKNPAIASATGRRVSRMIGDVA
VSTCVPAAPNVIVONSMRVSRRGTCSPLVSFRYEDGPILEGOLGSENRRLTR
DALEPCTVGHRRYFIFGGGYVYFBEVYSHQLSRADVTSTFTIDLITMLDEHPVP
LEVYTRHEIKDSGLDYTEVQRNQLHDLRFADIDTVIRADANAMFAGLCAFEFGNG
DLGRAVGVVGVGVAASVSSVSSFMSPFGALVGLVLAGLVAAFAFRVYLQ
ORNPMLALYPLTTKEIKTSDPGVGVGEGEGAGCGPDEAKLAEREMIRMALVSAM
ERTEHKKARKKGTALLSSKVTNNVLRKNKARYSPLEHDEAGDEDEL"

36. .101

/gene="UL27"

885

/gene="UL27"

/citation=[2]

/replace="t"

1221

/gene="UL27"

/citation=[2]

/replace="c"

1259

/gene="UL27"

/citation=[2]

/replace="t"

2282

/gene="UL27"

/citation=[2]

/replace="a"

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

REFERENCE

AUTHORS

TITLE

JOURNAL

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

PAGES

Alpha herpesvirinae; Simplexvirus.
1 (bases 1 to 2855)
Terhune, S.S. and Spear, P.G.
Variability of herpes simplex virus glycoproteins in clinical
specimens

Unpublished
2 (bases 274 to 2855)
Stuve, J.L., Brown-Shimer, S., Pachl, C., Najarian, R., Dina, D. and
Burke, R.L.

Structure and expression of the herpes simplex virus type 2
glycoprotein gB gene

J. Virol. 61 (2), 326-335 (1987)

3027364

3 (bases 1 to 2855)

Spear, P.G.

Direct Submission
Submitted (11-JUL-1994) Patricia G. Spear, Microbiology-Immunology,
Northwestern University Medical School, 303 East Chicago Avenue,
Chicago, IL 60611, USA

1. .2855
Location/Qualifiers

/organism="Human herpesvirus 2"

/mol_type="genomic DNA"

/strain="WTM1A"

/db_xref="taxon:10310"

36. .2750

/gene="UL27"

36. .2750

/note="gb: precursor peptide"

/product="glycoprotein B"

/protein_id="AA860546.1"

/db_xref="GI:522171"

/translation="MRGGSLICLVGALVAANAAPAPAPASGVAATYANG
GPASRPPVPSPATYKARKTKKPPKPPATPPDANATVAAGHATLRALREIKYE
NADAQFYVCPPTGATVVOEPORPCPTREGQYTGIAVFPENIAPYFKATMY
KDVTSQVMFGHRSQFMGI FEDRAVPFEEVIDIKINAGVRSATAYVRNMETAF
HRDDETDLKPAKVATRTSGMHTDLKNSRVAPFRYGTTCVIEVDARSV
YPRDEPLATGDPVYNSPYGREGSTHTSTYAADFKQVDFYARDLTAKATISB
TTRNLTLTPKFTVAMDVPPKPAVCTMTKQVDEMLRAEYGSFRSSDAISTFTT
NLTEYSLSRVDLGDICGRDAREALDRFARKYNATHIKVGPQYTLATGFLIAYOPL
LSNTLAEIYVREYREBODRKPRNATPAPLEAPASANSVERIKTSSIEFARLOFTYN
HIORHVDMLGRIAVANCELOHNETLTMEARKNPAIASATGRRVSRMIGDVA
VSTCVPAAPNVIVONSMRVSRRGTCSPLVSFRYEDGPILEGOLGSENRRLTR
DALEPCTVGHRRYFIFGGGYVYFBEVYSHQLSRADVTSTFTIDLITMLDEHPVP
LEVYTRHEIKDSGLDYTEVQRNQLHDLRFADIDTVIRADANAMFAGLCAFEFGNG
DLGRAVGVVGVGVAASVSSVSSFMSPFGALVGLVLAGLVAAFAFRVYLQ
ORNPMLALYPLTTKEIKTSDPGVGVGEGEGAGCGPDEAKLAEREMIRMALVSAM
ERTEHKKARKKGTALLSSKVTNNVLRKNKARYSPLEHDEAGDEDEL"

36. .101

/gene="UL27"

203

/gene="UL27"

/citation=[2]

/replace="c"

885

/gene="UL27"

/citation=[2]

/replace="t"

1221

/gene="UL27"

/citation=[2]

/replace="c"

1259

/gene="UL27"

/citation=[2]

/replace="a"

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

Alpha herpesvirinae; Simplexvirus.
1 (bases 1 to 2855)
Terhune, S.S. and Spear, P.G.
Variability of herpes simplex virus glycoproteins in clinical
specimens

Unpublished
2 (bases 274 to 2855)
Stuve, J.L., Brown-Shimer, S., Pachl, C., Najarian, R., Dina, D. and
Burke, R.L.

Structure and expression of the herpes simplex virus type 2
glycoprotein gB gene

J. Virol. 61 (2), 326-335 (1987)

3027364

3 (bases 1 to 2855)

Spear, P.G.

Direct Submission
Submitted (11-JUL-1994) Patricia G. Spear, Microbiology-Immunology,
Northwestern University Medical School, 303 East Chicago Avenue,
Chicago, IL 60611, USA

1. .2855
Location/Qualifiers

/organism="Human herpesvirus 2"

/mol_type="genomic DNA"

/strain="WTM1A"

/db_xref="taxon:10310"

36. .2750

/gene="UL27"

36. .2750

/note="gb: precursor peptide"

/product="glycoprotein B"

/protein_id="AA860546.1"

/db_xref="GI:522171"

/translation="MRGGSLICLVGALVAANAAPAPAPASGVAATYANG
GPASRPPVPSPATYKARKTKKPPKPPATPPDANATVAAGHATLRALREIKYE
NADAQFYVCPPTGATVVOEPORPCPTREGQYTGIAVFPENIAPYFKATMY
KDVTSQVMFGHRSQFMGI FEDRAVPFEEVIDIKINAGVRSATAYVRNMETAF
HRDDETDLKPAKVATRTSGMHTDLKNSRVAPFRYGTTCVIEVDARSV
YPRDEPLATGDPVYNSPYGREGSTHTSTYAADFKQVDFYARDLTAKATISB
TTRNLTLTPKFTVAMDVPPKPAVCTMTKQVDEMLRAEYGSFRSSDAISTFTT
NLTEYSLSRVDLGDICGRDAREALDRFARKYNATHIKVGPQYTLATGFLIAYOPL
LSNTLAEIYVREYREBODRKPRNATPAPLEAPASANSVERIKTSSIEFARLOFTYN
HIORHVDMLGRIAVANCELOHNETLTMEARKNPAIASATGRRVSRMIGDVA
VSTCVPAAPNVIVONSMRVSRRGTCSPLVSFRYEDGPILEGOLGSENRRLTR
DALEPCTVGHRRYFIFGGGYVYFBEVYSHQLSRADVTSTFTIDLITMLDEHPVP
LEVYTRHEIKDSGLDYTEVQRNQLHDLRFADIDTVIRADANAMFAGLCAFEFGNG
DLGRAVGVVGVGVAASVSSVSSFMSPFGALVGLVLAGLVAAFAFRVYLQ
ORNPMLALYPLTTKEIKTSDPGVGVGEGEGAGCGPDEAKLAEREMIRMALVSAM
ERTEHKKARKKGTALLSSKVTNNVLRKNKARYSPLEHDEAGDEDEL"

36. .101

/gene="UL27"

203

/gene="UL27"

/citation=[2]

/replace="c"

885

/gene="UL27"

/citation=[2]

/replace="t"

1221

/gene="UL27"

/citation=[2]

/replace="a"

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

967 c

936 g

433 t

520 a

Best Local Similarity 95.7%; Pred. No. 90;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAAACTCGTGCTCCAGCATG 28
|||||
Db 2076 CAAACTCGTGCTCCAGCATG 2054

RESULT 9
H52GB/c
LOCUS H52GB 3324 bp DNA linear VRL 15-MAY-1995
DEFINITION H52GB glycoprotein B gene, complete cds.
ACCESSION M14923
VERSION M14923.1 GI:330254
KEYWORDS glycoprotein; glycoprotein gb.
SOURCE Human herpesvirus 2
ORGANISM Human herpesvirus 2
VIRUSES; ddbDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.

REFERENCE 1
AUTHORS Bzik,D.J., Debroy,C., Fox,B.A., Pederson,N.E. and Person,S.
TITLE The nucleotide sequence of the gb glycoprotein gene of HSV-2 and comparison with the corresponding gene of HSV-1
JOURNAL Virology 155 (2), 322-333 (1986)
MEDLINE 87071654
PUBMED 3024391
COMMENT Original source text: Herpes simplex virus type 2 (strain HG52)

DNA
Draft entry and computer-readable sequence of (1) kindly provided by N.E. Pederson, 10-APR-1987.
The sequences of gb glycoprotein genes of HSV-1 and HSV-2 were compared and there was an overall amino acid homology of 85%. There were three major regions of homology at nucleotides 108-397, 484-785, and 835-873.
A membrane spanning region was found at position 2717-2923. The authors proposed that a signal cleavage-recognition site exists after nucleotide 613. This would give a mature peptide of 882 residues, from 614-3262.
Location/Qualifiers

FEATURES
SOURCE 1..3324
/organism="Human herpesvirus 2"
/mol_type="genomic DNA"
/strain="HG52"
/db_xref="taxon:10310"
/lab_host="HEL cell"
/note="ORF, putative"
/codon_start=1
/product="unknown protein"
/protein_id="AAA66439.1"
/db_xref="GI:807643"
/translation="VNCPLFDHSTHSPAPPTALYSVENVGLPHLKERLAPINGAGSGADMANVSEFQKFCYFDVSGITTPORAMKRIKLLIATILFAVYGCGELELRHDCRPTSEGLYRYPGVLYLYNSDCLVALIVESGPDGIGPRVVVYDVSILYSVLQHLAPRLAGSGDADP"
261..3322
/note="gb mRNA (5' end +/- 4 bp)"
548..3362
/codon_start=1
/product="glycoprotein gb"
/protein_id="AAA66440.1"
/db_xref="GI:330255"
/translation="MRGGGLICALVVALVAASAPAAPAPRAGSGVATVAANGGASRPVPPATTKARKRKTKPKPKPKATPPPDANAATVAAGATVRAIRIKVE NADAPFYCPPTGATVVOFEDPRACPTPRPESQNTTEIAYVEKENIAPYFKATMY KDVTSVQVPGHRSQFWGIFEDRAPVPEVIDINKAGCRSTAKTVRNMETTAR HRDDETMELMELPAKVAARTSGMHTTDLKNPSVVEAFHRTGTVNCIVEVARSV YPDEFLATGDFVMSPPYGRGSHETHTYADREKQYDGFYADLTAKASTP TTRNLITPKFTVANDVWPKPAVCTMTKQGVDEMLABAGSGSFRPSDAISTFT NUTETSLSRVLDGCTIGRDAREPIDRMPARKNAHTIKVGOPOVYQATGFLIAYOPL LSTNTALYVREYHEQRKRNATPAPLREKPSANAVERTKTSISIEPRLQPTTN VIOHVNMGLRIAVMCELOHETLTLNNEARKLPNIASTVGRRVASAMLDGVMA VATCVPAVDNVIYONSMRVSRSPTCYSRPLVSFRYEDGPLEGQLGENNDVLTTR

DALEPCTVGRGVFFPGGYVEEYAVSHQLSRADVTVSTPIDLITMLDEHFPV
LEVTTRHEIDSGILDTYEVORRNQDLHPADIDVIRADANAMAPGACAPFEGMG
DLGAVGKVMGVGVSAVSGVSPNSNPPGALVGLVAGLVAFAPFRVYQL
ORNMKALYPLLTKEKLTSDPVGGEAGEGSGDEAKLAEAREMIRYMALVSAM
ERTHEKARKGTGALLSKTYTNMVLRRNARYSPLINDEBAGEDEDL"
BASE COUNT 600 a 1151 c 1063 g 510 t
ORIGIN 242 bp upstream of NruI site; between 0.370 and 0.348 map units.

Query Match 68.5%; Score 22.6; DB 14; Length 3324;
Best Local Similarity 95.7%; Pred. No. 89;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAAACTCGTGCTCCAGCATG 28
|||||
Db 2588 CAAACTCGTGCTCCAGCATG 2566

RESULT 10
BD165792/c
LOCUS BD165792 3472 bp DNA linear PAT 17-JAN-2003
DEFINITION Vaccine for use in the therapeutic treatment of HSV.
ACCESSION BD165792
VERSION BD165792.1 GI:27871604
KEYWORDS JP 2002167398-A/1.
SOURCE Herpes simplex virus 7
ORGANISM Herpes simplex virus 7
VIRUSES; ddbDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.

REFERENCE 1
AUTHORS Burke,R.L., Pachl,C. and Valenzuela,P.D.T.
TITLE Vaccine for use in the therapeutic treatment of HSV
JOURNAL Patent: JP 2002167398-A 1 11-JUN-2002;
CHIRON CORP
OS Herpes simplex virus 7
BN JP 2002167398-A/1
PD 11-JUN-2002
PF 12-SEP-2001 JP 2001277340
PR 20-OCT-1986 US 921,213
PI RAB LYN BURKE, CAROL PACHL, PABLO D T VALENZUELA PC
COTK14/035, A61K39/245, A61P31/22, C12N15/09, C12N15/00, C12N15/00
CC Vaccine for use in the therapeutic treatment of HSV FH Key

FEATURES
SOURCE 1..3472
/organism="Herpes simplex virus 7"
/mol_type="genomic DNA"
/db_xref="taxon:57277"
FT CDS Location/Qualifiers (309)..(3023).

BASE COUNT 639 a 1202 c 1069 g 562 t
ORIGIN

Query Match 68.5%; Score 22.6; DB 6; Length 3472;
Best Local Similarity 95.7%; Pred. No. 89;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAAACTCGTGCTCCAGCATG 28
|||||
Db 2349 CAAACTCGTGCTCCAGCATG 2327

RESULT 11
108685/c
LOCUS 108685 3472 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent WO 8802634.
ACCESSION 108685
VERSION 108685.1 GI:588605
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3472)
AUTHORS Burke,R.L., Pachl,C. and Valenzuela,P.D.T.

BASE COUNT	ORIGIN
639 a	3 bp upstream of NruI site; 0.344 to 0.367 mu on HSV-2 map.
1202 c	
1069 g	
562 t	

VTYSQWFGHRI5QFMGI FEDRAPVPEEVIIDKJNAGVC8R3AKTYVRBNMNETAFHRR
 DDHETDMLRPAKATRSRGMHTTDLKYN5SREARHGTTCNCIVEEIVDASVPP
 YDEVLVALTGDFVNMSPYGRBESGHTTYADPRQNDVGFYATTCITABARASPTT
 RNLLTTRKFTVAMDWPKRPVACGTMTEMOEDMLRREYGG5FNFSDA15TFTYTNL

Query Match	68.5%	Score 22.6	DB 14	Length 3715
Beat Local Similarity	95.7%	Pred. No. 89		
Matches 22	Conservative 1	Mismatches 0	Indels 0	Gaps 0
BASE COUNT	683 a	1287 c	1135 g	610 t
QUERY	<p>6 CAAATCGTGTCTCCGACGATG 28</p> <p> </p> <p>2590 CAAATCGTGTCTCCGACGATG 2568</p>			
LOCUS	HSV2HG52	154746 bp	DNA	linear
DEFINITION	Herpes simplex virus type 2 (strain HG52), complete genome.			
ACCESSION	286039			
VERSION	286039.2	GI:6572414		
KEYWORDS	<p>capsid associated tegument protein; capsid protein; complete genome; deoxyribonuclease; deoxyribonuclease; DNA polymerase; envelope protein; glycoprotein B; glycoprotein C; glycoprotein D; glycoprotein E; glycoprotein G; glycoprotein H; glycoprotein I; glycoprotein J; glycoprotein K; glycoprotein L; glycoprotein M; host shut-off factor; integral membrane protein; major capsid protein; membrane-associated phosphoprotein; minor capsid protein; neurovirulence factor; protease; protein kinase; ribonucleotide reductase large subunit; ribonucleotide reductase small subunit; R1 gene; R2 gene; R3 gene; tegument protein; thymidine kinase; UL1 gene; UL10 gene; UL11 gene; UL12 gene; UL13 gene; UL14 gene; UL15 gene; UL16 gene; UL17 gene; UL18 gene; UL19 gene; UL20 gene; UL21 gene; UL22 gene; UL23 gene; UL24 gene; UL25 gene; UL26 gene; UL26.5 gene; UL27 gene; UL28 gene; UL29 gene; UL30 gene; UL31 gene; UL32 gene; UL33 gene; UL34 gene; UL35 gene; UL36 gene; UL37 gene; UL38 gene; UL39 gene; UL40 gene; UL41 gene; UL42 gene; UL43 gene; UL44 gene; UL45 gene; UL46 gene; UL47 gene; UL48 gene; UL49 gene; UL49A gene; UL50 gene; UL51 gene; UL52 gene; UL53 gene; UL54 gene; UL55 gene; UL56 gene; UL6 gene; UL7 gene; UL8 gene; UL9 gene; uracil-DNA glycosylase; US1 gene; US10 gene; US11 gene; US12 gene; US2 gene; US3 gene; US4 gene; US5 gene; US6 gene; US7 gene; US8 gene; US9A gene; US9 gene.</p>			
SOURCE	Human herpesvirus 2			
ORGANISM	Human herpesvirus 2			
REFERENCE	<p>1 Alpha herpesvirinae; Simplexvirus.</p> <p>2 (bases 1 to 154746)</p> <p>McGeoch, D. J., Moss, H. M., McNab, D. and Pringle, M. C.</p>			
AUTHORS	DNA sequence and genetic content of the HindIII 1 region in the short unique component of the herpes simplex virus type 2 genome: identification of the gene encoding glycoprotein G, and evolutionary comparisons			
JOURNAL	J. Gen. Virol. 68 (Pt 1), 19-38 (1987)			
MEDLINE	87111457			
PIRBASE	3027242			
REFERENCES	2 (bases 1 to 154746)			
7 AUTHORS	Everett, R. D. and Fenwick, M. L.			

FEATURES	COMMENT
JOURNAL	Comparative DNA sequence analysis of the host shutoff genes of different strains of herpes simplex virus: type 2 strain HG52
MEDLINE	encodes a truncated UL41 product
PUBMED	J. Gen. Virol. 71 (Pt 6), 1387-1390 (1990)
REFERENCE	3 (bases 1 to 154746)
AUTHORS	McGeoch,D.J., Cunningham,C., McIntyre,G. and Dolan,A.
TITLE	Comparative sequence analysis of the long repeat regions and adjoining parts of the long unique regions in the genomes of herpes simplex viruses types 1 and 2
JOURNAL	J. Gen. Virol. 72 (Pt 12), 3057-3075 (1991)
MEDLINE	92113549
PUBMED	1662697
REFERENCE	4 (bases 1 to 154746)
AUTHORS	Barnett,B.C., Dolan,A., Telford,E.A., Davison,A.J. and McGeoch,D.J.
TITLE	A novel herpes simplex virus gene (UL49A) encodes a putative membrane protein with counterparts in other herpesviruses
JOURNAL	J. Gen. Virol. 73 (Pt 8), 2167-2171 (1992)
MEDLINE	92356101
PUBMED	1322965
REFERENCE	5
AUTHORS	Dolan,A.
TITLE	Direct Submission
JOURNAL	Submitted (25-FEB-1997) A. Dolan, MRC Virology Unit, Church Street, Glasgow, G11 5UR, UK
REMARK	revised bu [6]
REFERENCE	6 (bases 1 to 154746)
AUTHORS	Dolan,A.
TITLE	Direct Submission
JOURNAL	Submitted (08-DEC-1999) A. Dolan, MRC Virology Unit, Church Street, Glasgow, G11 5UR, UK
COMMENT	On Dec 13, 1999 this sequence version replaced gi:1869820.
FEATURES	Location/Qualifiers
source	1. .154746
	/organism="Human herpesvirus 2"
	/mol_type="genomic DNA"
	/strain="HG52"
	/db_xref="taxon:10310"
	1. .254
	/note="a sequence"
	255
	/note="start of TRL (terminal copy of Long Repeat region) "
	440..1379
	/gene="RL1"
	join(440..934,1089..1379)
	/gene="RL1"
	/codon_start=1
	/protein_id="CAB06759.1"
	/db_xref="GI:1869821"
	/db_xref="SWISS-PROT:P28283"
	/translation="MSRRGPRRRGPRRRPGCAVAPRPQAPVPPGALPTADSQAVPAPDSTGAVASAPAASTLRMLTVQADSDADYAGNDADAMNSPSEGGKAPAEAPAAAPACPPPRPKERGGPQRPILPPIALRTTTLVRLARSRRPPAPPADAPKRCVPSFPRVYRHLVANETARKLRARRRSWABERADRRFRRVAAAEAVIGPCLEPEARAKRARARAHEDGPAEEEAALAGSSSAAAGPRAV"
	440..934
	/gene="RL1"
	/number=1
	446..460
	/rpt_family="7"
	461..547
	/rpt_family="6"
	943..1070
	/rpt_family="5"
	1089..1379
	/gene="RL1"
	/product="neurovirulence factor"
	/number=2
	1738..1743
	2303..5368
	/gene="RL2"
	join(2303..2377,2785..3462,3644..5368)

/gene="RL2"
 /codon_start=1
 /protein_id="CAB06760.1"
 /db_xref="GI:1869822"
 /db_xref="SWISS-PROT:p28284"
 /translation="MEPRGTSRADPQPERPRPQPTGTOPAPAHGMLNDMOWLAS
 SISEETETGIDDDLDHSDTSEAGSTPTLMEBAIADAPAPAPRERGSPAPAPAS
 AOGSGEYVGEERBAAGGGGVCAVCTDELAPELRCSFCLHPRCIPCKMTIPLME
 NTPCANTPVALLVGYTVASSGSFTIPIYNDPRIVYEAAVRACTAVDTPTGNERPT
 APRSLSLGHTVRLALSTPTPEWGTDEDDDLADVIVPAPAPRARRGGAGATRGTA
 SQPARIAPAPAPAPSSSSGGAFLRAGVSGGGGPAAVAVPRVASLPPAGGGRAA
 ARRVEDAAEAAGRTPPAROPPAOPEPVIISDPSPPRRPAGGPGLSFVSSSAQV
 SSPGGGGLPGSSGGAAPRAVAVRVSPPRAAAPVVSADAAAGAPAVPVDADA
 RABPRMTOAQDTTOAGLGRAGATDARGSGGPGAGGPGVPRGCTNTPAAPHAGGE
 AARPRKRGSDSGPAASSASSASSAAPRPEPLAQYGAARRAPRAVDSDSGRGHP
 APASGAAPPSAPSSQAALVAASSSSASSASSSSASSASSASSASSASSASSSSG
 SSSAGAGGVSASAGAGRRRTSLGPPAAAPRGRCARKTRHAEGGEGEGADPAAP
 GLTRRLPLAGVSSVALAPVYMKVTGDCLPVLDMEHTIGAYVVLVDTGTNVDLML
 AAAPMSRRLTPEHARRCPRPDYPPTPPASBWSLMTPTPGNMLFDGGLVGLDFFH
 GLRSHRPSREGCAPAPAGDAPAHGGE"
 2303..2377
 /gene="RL2"
 /number=1
 2785..3462
 /gene="RL2"
 /number=2
 3468..3607
 /rpt_family="4"
 3644..5368
 /gene="RL2"
 /number=3
 /function="modulator of cell state and gene expression"
 4655..4771
 /rpt_family="3"
 complement(4811)
 /note="proposed LAR splice acceptor site"
 5618..5623
 complement(7025)
 /note="proposed LAR splice donor site (5' end of stable
 LAR species)"
 complement(7732)
 /note="primary transcription initiation site for latency
 associated transcripts (LAR's)"
 9047..9193
 /rpt_family="2"
 9276..9297
 /rpt_family="1"
 9298
 /note="start of UL (long Unique region)"
 9427..10101
 /gene="UL1"
 9427..10101
 /gene="UL1"
 /codon_start=1
 /product="viralion glycoprotein L"
 /protein_id="CAB06761.1"
 /db_xref="GI:1869823"
 /db_xref="SWISS-PROT:p28278"
 /translation="MGFVCLRGVYMGMGMGMSQATETYLRSVIAEYVDILVPE
 NRTPADVSMKTEAPSVLDIYARIDGFIETPHCPGIDLTLMDDHARALVNPFLPAAE
 FLEIDISHVFPADTQETTRALVYEIRDALGSRQAAVSHA.PVAGCIVPYSRRRR
 VGRRLIRPANTTSTWEPVPSDDDEASSQSKPLATQPPVLAISNAPRRHVSPTGRRRRR
 TRLRRN"
 10211..10978
 /gene="UL2"
 10211..10978
 /gene="UL2"
 /codon_start=1
 /product="uracil-DNA glycosylase"
 /protein_id="CAB06762.1"
 /db_xref="GI:1869824"
 /db_xref="SWISS-PROT:p28275"
 /translation="MFSASTPEOPGLGSGDATPELPVPLDWAARFRAFLIDDAWMA

[illegible]

LAETVGRTERHQSPPGDFPNLKRMQDLQAQDFSKFQPLNSKLLLEVVDMLANDIAQ
LMLVROEESORPIOMVKGAFEGTIGHGPGHGYGEGAGESIDDAEMVAVARDKPMYDS
IFVTLSPVDSKITGANAKKENVRSKLPNSVIGIKWKLADIDKQGLDDDEFALANHLI
KVXLEGHELPHNELPAHLLPRPRGKLPSPDGVGSHSDGC"

BASE COUNT 541 a 639 c 653 g 448 t
ORIGIN

Query Match 67.9%; Score 22.4; DB 9; Length 2281;
Best Local Similarity 88.5%; Pred. No. 1.1e+02;
Matches 23; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 4 TGCAACTCGTGTCTCCAGCATGC 29
|||||
|||

Db 1550 TGCAACTCGTGTCTCCAGCATGC 1525
|||||
|||

Search completed: September 16, 2003, 19:14:19
Job time : 724.713 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 17:02:34 ; Search time 165.327 Seconds
(without alignments)
538.821 Million cell updates/sec

Title: US-09-594-065-4

Perfect score: 33 ccctgcaactcgttctccacgacgacgagcggg 33

Sequence: 1

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

5105512

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22.6	68.5	2645	19	AAV62158
2	22.6	68.5	2724	8	AAV62158
3	22.6	68.5	2727	6	AAV62158
4	22.6	68.5	2841	19	AAV62158
5	22.6	68.5	3471	18	AAV62158
6	22.6	68.5	3472	9	AAV62158
7	22.6	68.5	3472	14	AAV62158
8	22.6	68.5	117213	19	AAV62158

9	22.6	68.5	154746	24	AAV62158
10	22.4	67.9	2850	22	AAV62158
11	22.4	67.9	2850	22	AAV62158
12	22.4	67.9	3327	22	AAV62158
13	22.4	64.8	1494	25	AAV62158
14	21.4	64.8	2088	12	AAV62158
15	21.4	64.8	2712	8	AAV62158
16	21.4	64.8	2925	12	AAV62158
17	21.4	64.8	3098	7	AAV62158
18	21.4	64.8	3461	8	AAV62158
19	21.4	64.8	3465	12	AAV62158
20	21.4	64.8	3472	14	AAV62158
21	21.4	64.8	3473	18	AAV62158
22	21.4	64.8	3474	9	AAV62158
23	21.4	64.8	3642	6	AAV62158
24	21.4	64.8	3996	6	AAV62158
25	21.4	64.8	69936	21	AAV62158
26	21.4	64.8	349980	21	AAV62158
27	21.4	64.8	1437668	21	AAV62158
28	20.8	63.0	570	23	AAV62158
29	20.8	63.0	764	22	AAV62158
30	20.8	63.0	1549	22	AAV62158
31	20.8	63.0	2121	22	AAV62158
32	20.8	63.0	2240	22	AAV62158
33	20.8	63.0	2240	22	AAV62158
34	20.8	63.0	3437	22	AAV62158
35	20.2	61.2	2339	17	AAV62158
36	20.2	61.2	2339	17	AAV62158
37	20.2	61.2	3059	24	AAV62158
38	20.2	61.2	3138	21	AAV62158
39	20.2	61.2	3138	24	AAV62158
40	20.2	61.2	3138	24	AAV62158
41	19.8	60.0	2698	23	AAV62158
42	19.8	60.0	3679	22	AAV62158
43	19.8	60.0	3679	22	AAV62158
44	19.8	60.0	3679	25	AAV62158
45	19.8	60.0	3679	25	AAV62158

ALIGNMENTS

RESULT 1	AAV62158	standard; DNA; 2645 BP.
AAV62158		
AC	AAV62158;	
DT	18-DEC-1998	(first entry)
XX		
DE	HSV-2 strain SBS Contig ID 14 DNA sequence.	
XX		
KW	HSV-2 strain SBS; immunological response induction; therapy;	
KM	antiviral identification; viral protein inhibitor; ss.	
XX		
OS	Herpes simplex virus type 2.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	complement (81..2645)
FT		/*tag= a
FT		/product= "ORF#1 protein"
XX		
PN	NO9820016-A1.	
PD	14-MAY-1998.	
XX		
PF	31-OCT-1997;	97MO-US20016.
XX		
PR	09-JUN-1997;	97US-0049018.
XX		
PR	04-NOV-1996;	96US-0030279.
XX		
PA	(SMIR) SMITHKLINE BEECHAM CORP.	
XX		

Human herpesvirus
Human polynucleoti
CDNA of human EBD2
Human polynucleoti
N. gonorrhoeae nuc
Truncated HSVB ge
Herpes Simplex Vir
HSV surface antige
Herpes simplex vir
Recombinant herpes
HSV8 gene. Herpe
Glycoprotein B (gB
Herpes simplex vir
Sequence of Herpes
DNA fragment encod
Sequence encoding
N. meningitidis pa
Neisseria meningit
N. meningitidis B
DNA encoding novel
Human nervous syst
Human nervous syst
Human nervous syst
Human nervous syst
Human nervous syst
Nucleic acid sequ
Interleukin bindin
Interleukin bindin
Breast carcinoma r
Human RING finger
Human cDNA encodin
Human breast cance
Drosophila melanog
Protein PRO293 CDN
Human PRO293 CDNA.
Novel human secret
CDNA encoding huma

PI Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB,
PI Esser KM, Leary JJ;
XX
XX MPI, 1998-286847/25.
DR P-PSDB; AAM72113.
XX
XX Herpes simplex virus type-2 sequences - useful in, e.g. prevention
PT and treatment of infection or inducing immunological response in
PT mammal
XX
XX
PS Claim 1; Page 406-407; 748pp; English.
XX
XX This sequence represents a Herpes simplex virus type-2 (HSV-2) DNA
CC sequence of the invention. This sequence was isolated from HSV-2 strain
CC SBS (deposited as ATCC VR-2546), is designated Contig ID 14, and encodes
CC a HSV-2 protein. The protein can be used for the treatment or
CC prevention of disease, to induce an immunological response in a mammal or
CC to identify inhibitors, activators or novel antivirals. Antagonists of
CC the proteins can be used to inhibit a viral polypeptide. The DNA sequence
CC or a vector containing it can also be used to induce an immunological
CC response in a mammal.
XX
SQ Sequence 2645 BP; 404 A; 837 C; 892 G; 512 T; 0 other;

Query Match 68.5%; Score 22.6; DB 19; Length 2645;
Best Local Similarity 95.7%; Pred. No. 17;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAAACTCGTGTCTCCGACGATG 28
DB 755 CAAACTCGTGTCTCCGACGATG 777

RESULT 2
ID AAN71399/C
XX AAN71399 standard; DNA; 2724 BP.
XX
AC AAN71399;
XX
XX 25-MAR-2003 (updated)
DT 30-APR-1991 (first entry)
XX
XX Herpes Simplex Virus-2 gB from p52BXX.
DE
XX
XX Vaccine; prophylaxis; treatment; Herpes Simplex Virus-1;
KW glycoprotein; gB; ss.
XX
XX Herpes simplex virus type 2 HG52.
OS
XX
XX US4642333-A.
PN
XX
XX 10-FEB-1987.
PD
XX
XX 20-JUN-1984; 84US-0622496.
PE
XX
XX 20-JUN-1984; 84US-0622496.
PR
XX
XX (PERS/) PERSON S.
PA
XX
XX
PI Person S;
XX
XX MPI, 1987-056354/08.
DR P-PSDB; AAP71136.
XX
XX Amino acid chain of glyco:protein B of HSV-1 and 2 - prepd. as
PT recombinant and used for vaccines for herpes simplex virus types 1
PT and 2.
XX
XX
XX Example; Table 2; 16pp; English.
PS
XX
XX 'N' indicated nucleotide to be determined (corresp. to residues
CC 17-41 of HSV-1 gB).
CC
QC A pure non-glycosylated amino acid (AA) chain comprising a sequence

CC corresponding to that occurring in glycoprotein B of HSV-1 or HSV-2
CC virus which is antigenic to HSV-1 of HSV-2, which contains no more
CC than 750 AA residues, and which includes AA residues 135-649
CC inclusive is claimed. It can be used to produce vaccines for
CC prophylaxis and treatment of HSV-1 and HSV-2.
CC See AAN71303 for the HSV-1 sequence.
CC (Updated on 25-MAR-2003 to correct PR field.)
XX
XX
SQ Sequence 2724 BP; 494 A; 911 C; 854 G; 390 T; 75 other;

Query Match 68.5%; Score 22.6; DB 8; Length 2724;
Best Local Similarity 95.7%; Pred. No. 18;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAAACTCGTGTCTCCGACGATG 28
DB 2050 CAAACTCGTGTCTCCGACGATG 2028

RESULT 3
ID AAN50517/C
XX AAN50517 standard; DNA; 2727 BP.
XX
AC AAN50517;
XX
XX 21-NOV-1991 (first entry)
DT
XX
XX Sequence encoding the glycoprotein B (gB) of Herpes simplex virus
DE type 2 (HSV-2).
DE
XX
XX Vaccine; glycoprotein B; ss.
KW
XX
XX Herpes simplex virus type 2;
OS
XX
XX Key Location/Qualifiers
FH 1..2727
FT CDS /*tag= a
FT
XX
XX EP133063-A.
PN
XX
XX 13-FEB-1985.
PD
XX
XX 22-JUN-1984; 84EP-0401312.
PE
XX
XX 16-SEP-1983; 83US-0532996.
PR 23-JUN-1983; 83US-0506986.
PR 20-JUN-1984; 84US-0622496.
XX
XX (PERS/) PERSON S.
PA
XX
XX
PI Person S;
XX
XX MPI, 1985-039636/07.
DR
XX
XX New amino acid chains of glyco:protein B of herpes simplex virus
PT - prepd. by recombination DNA methods for use in vaccines
PT
XX
XX Disclosure; Table 2, Page 35-35C; 40pp; English.
PS
XX
XX The inventors claim a non-glycosylated amino acid chain comprising a
CC sequence corresponding to that occurring in glycoprotein B of HSV-1
CC or HSV-2. Preferably, when the chain comprises a sequence corresp.
CC to that occurring in glycoprotein B of HSV-2, it contains 4-750 AA
CC residues. The chain esp. has a molecular wt. of 65000 daltons and
CC includes AA residues 155-629 from the HSV-2 glycoprotein B sequence.
XX
XX
SQ Sequence 2727 BP; 493 A; 913 C; 855 G; 391 T; 75 other;

Query Match 68.5%; Score 22.6; DB 6; Length 2727;
Best Local Similarity 95.7%; Pred. No. 18;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAAACTCGTGTCTCCGACGATG 28

Db 2053 CAAACTGCTGCTCCTCCAGCATG 2031

RESULT 4

AAV62146/C
ID AAV62146 standard; DNA; 2841 BP.

AAV62146;

07-DEC-1998 (first entry)

HSV-2 strain SBS Contig ID 91 DNA sequence.

HSV-2 strain SBS; immunological response induction; therapy;
antiviral identification; viral protein inhibitor; ss.

Herpes simplex virus type 2.

Key Location/Qualifiers
CDS 358..2745
/*tag= a

MO9820016-A1.

14-MAY-1998.

31-OCT-1997; 97WO-US20016.

09-JUN-1997; 97US-0049018.

04-NOV-1996; 96US-0030279.

(SMIK) SMITHKLINE BEECHAM CORP.

Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB;

Besser KM, Leary JO;

WPI; 1998-286847/25.

P-PSDB; AAW72062.

Herpes simplex virus type-2 sequences - useful in, e.g. prevention

and treatment of infection or inducing immunological response in

mammal

Claim 1; Page 284-285; 748bp; English.

This sequence represents a Herpes simplex virus type-2 (HSV-2) DNA

sequence of the invention. This sequence was isolated from HSV-2 strain

SBS (deposited as ATCC VR-2546), is designated Contig ID 91, and encodes

a HSV-2 protein. The proteins can be used for the treatment or

prevention of disease, to induce an immunological response in a mammal or

to identify inhibitors, activators or novel antivirals. Antagonists of

the proteins can be used to inhibit a viral polypeptide. The DNA sequence

or a vector containing it can also be used to induce an immunological

response in a mammal.

Sequence 2841 BP; 577 A; 921 C; 915 G; 425 T; 3 other;

Query Match 68.5%; Score 22.6; DB 19; Length 2841;

Best Local Similarity 95.7%; Pred. No. 18;

Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DT 26-FEB-1998 (first entry)

XX Herpes simplex virus type 1 glycoprotein B encoding DNA.

XX Herpes simplex virus; HSV; type I; type II; glycoprotein B; vaccine;

XX membrane protein; virus-specific glycoprotein;

XX transmembrane anchor region; ss.

XX Herpes simplex virus type 1.

Key Location/Qualifiers

CDS 55..62

/*tag= a

/*note= "putative"

TATA_signal 125..131

/*tag= b

/*note= "putative"

CDS 309..3023

/*tag= c

/*tag= d

/*note= "putative"

polyA_signal 3409..3416

/*tag= e

/*note= "putative"

US5648079-A.

15-JUL-1997.

08-DEC-1994; 94US-0351875.

20-OCT-1986; 86US-0921730.

06-APR-1984; 84US-0597284.

20-SEP-1990; 90US-0587179.

21-DEC-1992; 92US-0933415.

18-OCT-1993; 93US-0138717.

(CHIR) CHIRON CORP.

Burke RL, Pachl C, Valenzuela PDT;

WPI; 1997-372022/34.

P-PSDB; AAW34552.

Vaccines against herpes simplex virus infection - containing

recombinant HSV glycoprotein B

Disclosure; Fig 4A-F; 33pp; English.

The present sequence is from the Herpes simplex virus (HSV) type I

strain Paton and encodes a glycoprotein B. HSV is a double stranded

virus packaged within an icosahedral nucleocapsid enveloped within a

membrane. The membrane includes a number of virus-specific

glycoproteins, with glycoprotein B being one of the most abundant.

Glycoprotein B from both HSV type I and type II are cross reactive. New

vaccines comprising recombinantly produced glycosylated glycoprotein B

that has a deletion of at least part of the transmembrane anchor region,

in combination with a carrier and an adjuvant have been produced. The

vaccines are used for immunising humans against HSV (HSV type 1 or

HSV type 2) by vaccination before or after primary infection with HSV.

Sequence 3471 BP; 639 A; 1201 C; 1069 G; 562 T; 0 other;

Query Match 68.5%; Score 22.6; DB 18; Length 3471;

Best Local Similarity 95.7%; Pred. No. 18;

Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

6 CAAACTGCTGCTCCTCCAGCATG 28

AAAT93650/C

AAAT93650 standard; DNA; 3471 BP.

AAAT93650;

```

RESULT 6
ID AAN80907/C standard; DNA; 3472 BP.
XX
AC AAN80907;
XX
DT 25-MAR-2003 (updated)
DT 09-MAR-1992 (first entry)
XX
DE Sequence of Herpes simplex virus (HSV) glycoprotein B (gB) 1 DNA.
XX
KM Vaccine; herpes simplex virus; therapy; ss.
XX
OS Herpes simplex virus.
XX
FH Key Location/Qualifiers
FT CDS 309..3023
FT /*tag= a
XX
PA MO8802634-A.
XX
PI 21-APR-1986.
XX
PF 20-OCT-1987; 87WO-US02709.
XX
PR 20-JUL-1987; 87US-0079605.
PR 20-OCT-1986; 86US-0921213.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Burke LR, Pachl C, Valenzuela P;
XX
DR WPI; 1988-119368/17.
XX
P-PSDB; AAP80914.
XX
DE Vaccine for treatment of herpes simplex virus - contains
PT recombinant HSV glyco:protein B and D
XX
PS Example; Fig 4; 71pp; English.
XX
CC Prep'n. of recombinant gB and gD is described in WO88504587. The
CC amino acid sequence and DNA sequence for gB 1 presented in AAP80914
CC and AAN80907 differ from that originally presented in Table 1 of
CC International Publication No. WO 85/04587. The DNA sequence in Table
CC 1 contains an error in that an additional nucleotide (G) is listed
CC at position 607 which resulted in a shift in reading frame relative
CC to AAN80907 from which this nucleotide has been deleted.
CC (Updated on 25-MAR-2003 to correct PA field.)
CC
SQ Sequence 3472 BP; 639 A; 1202 C; 1068 G; 563 T; 0 other;
XX
Query Match 68.5%; Score 22.6; DB 9; Length 3472;
Best Local Similarity 95.7%; Pred. No. 18;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
CY 6 CAAACTCGTGKTCCTCCAGCATG 28
DB 2349 CAAACTCGTGKTCCTCCAGCATG 2327
XX
RESULT 7
AAQ48496/C
ID AAQ48496 standard; DNA; 3472 BP.
XX
AC AAQ48496;
XX
DT 25-MAR-2003 (updated)
DT 29-MAR-1994 (first entry)
XX
DE Glycoprotein B (gB1).
XX
KM Glycoprotein; gB1; gB2; herpes simplex virus; HSV-1; HSV-2;
KM flanking; initiation; termination; transcription; translation;

```

```

KM regulatory sequence; ss.
XX
OS Herpes simplex virus 1 strain Patton.
XX
FH Key Location/Qualifiers
FT CDS 309..3023
FT /*tag= a
XX
PA US5244792-A.
XX
PI 14-SEP-1993.
XX
PF 20-SEP-1990; 90US-0587179.
XX
PR 06-APR-1984; 84US-0597784.
PR 20-OCT-1986; 86US-0921730.
PR 20-SEP-1990; 90US-0587179.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Burke RL, Pachl C, Valenzuela PDT;
XX
DR WPI; 1993-302641/38.
XX
P-PSDB; AAK41778.
XX
DE DNA construct for expressing HSV glycosylated polypeptide -
PT useful for vaccinating against HSV-1 and -2 infections
XX
PS Disclosure; Fig 4; 33pp; English.
XX
CC DNA constructs for expressing a glycosylated polypeptide in a
CC eukaryotic host cell comprises (1) an oligonucleotide sequence (OS)
CC free of natural flanking sequences, encoding glycoprotein B (gB)
CC of HSV, or its fragments, and (2) 5' initiating and 3' terminating
CC transcriptional and translational regulatory sequences flanking OS,
CC at least one of these sequences not being from HSV. The HSV-1 and
CC HSV-2 sequences are given in AAQ48496 and AAQ48497 respectively.
CC (Updated on 25-MAR-2003 to correct PF field.)
CC
SQ Sequence 3472 BP; 639 A; 1202 C; 1069 G; 562 T; 0 other;
XX
Query Match 68.5%; Score 22.6; DB 14; Length 3472;
Best Local Similarity 95.7%; Pred. No. 18;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
CY 6 CAAACTCGTGKTCCTCCAGCATG 28
DB 2349 CAAACTCGTGKTCCTCCAGCATG 2327
XX
RESULT 8
AAV62176
ID AAV62176 standard; DNA; 117213 BP.
XX
AC AAV62176;
XX
DT 13-JAN-1999 (first entry)
XX
DE HSV-2 strain SBS Contig ID 15 DNA sequence.
XX
KM HSV-2 strain SBS; immunological response induction; therapy;
KM antiviral identification; viral protein inhibitor; ss.
XX
OS Herpes simplex virus type 2.
XX
FH Key Location/Qualifiers
FT CDS 755..1297
FT /*tag= a
FT /product= "ORF#1 protein"
FT /note= "encoded protein shown in AAV72170"
FT 1170..2174
FT /*tag= b
FT /product= "ORF#2 protein"

```

FT /note= "encoded protein shown in AAW72171"
 FT 2229..2930
 FT /*tag= c
 FT /product= "ORF#3 protein"
 FT /note= "encoded protein shown in AAW72172"
 FT complement (3130..3735)
 FT /*tag= d
 FT /product= "ORF#4 protein"
 FT /note= "encoded protein shown in AAW72173"
 FT complement (3802..6447)
 FT /*tag= e
 FT /product= "ORF#5 protein"
 FT /note= "encoded protein shown in AAW72174"
 FT 6017..8482
 FT /*tag= f
 FT /product= "ORF#6 protein"
 FT /note= "encoded protein shown in AAW72250"
 FT 6026..8482
 FT /*tag= g
 FT /product= "ORF#6f protein"
 FT /note= "encoded protein shown in AAW72249"
 FT 6065..8482
 FT /*tag= h
 FT /product= "ORF#6e protein"
 FT /note= "encoded protein shown in AAW72248"
 FT 6167..8482
 FT /*tag= i
 FT /product= "ORF#6d protein"
 FT /note= "encoded protein shown in AAW72247"
 FT 6296..8482
 FT /*tag= j
 FT /product= "ORF#6c protein"
 FT /note= "encoded protein shown in AAW72246"
 FT 6326..8482
 FT /*tag= k
 FT /product= "ORF#6b protein"
 FT /note= "encoded protein shown in AAW72245"
 FT 6446..8482
 FT /*tag= l
 FT /product= "ORF#6a protein"
 FT /transl_except= (pos: 7400..7402, aa: Ala-Ala)
 FT /transl_except= (pos: 7481..7486, aa: Ile)
 FT /note= "encoded protein shown in AAW72244"
 FT 8457..9347
 FT /*tag= m
 FT /product= "ORF#7 protein"
 FT /note= "encoded protein shown in AAW72175"
 FT complement (9604..11855)
 FT /*tag= n
 FT /product= "ORF#8 protein"
 FT /transl_except= (pos: 11635..11636, aa: Ala)
 FT /note= "encoded protein shown in AAW72176"
 FT complement (11905..14508)
 FT /*tag= o
 FT /product= "ORF#9b protein"
 FT /note= "encoded protein shown in AAW72222"
 FT complement (11905..14520)
 FT /*tag= p
 FT /product= "ORF#9a protein"
 FT /note= "encoded protein shown in AAW72223"
 FT 14399..15802
 FT /*tag= q
 FT /product= "ORF#10 protein"
 FT /note= "encoded protein shown in AAW72177"
 FT complement (15996..16286)
 FT /*tag= r
 FT /product= "ORF#11 protein"
 FT /note= "encoded protein shown in AAW72178"
 FT complement (16202..18064)
 FT /*tag= s
 FT /product= "ORF#12 protein"
 FT /note= "encoded protein shown in AAW72179"
 FT complement (18105..19661)

FT /*tag= t
 FT /product= "ORF#13 protein"
 FT /note= "encoded protein shown in AAW72180"
 FT complement (19415..20074)
 FT /*tag= u
 FT /product= "ORF#14 protein"
 FT /note= "encoded protein shown in AAW72181"
 FT 20155..21453
 FT /*tag= v
 FT /product= "ORF#15 protein"
 FT /note= "encoded protein shown in AAW72182"
 FT complement (21326..42291)
 FT /*tag= w
 FT /product= "ORF#16 protein"
 FT /note= "encoded protein shown in AAW72183"
 FT complement (22546..24654)
 FT /*tag= x
 FT /product= "ORF#17 protein"
 FT /note= "encoded protein shown in AAW72184"
 FT 24684..25955
 FT /*tag= y
 FT /product= "ORF#18 protein"
 FT /note= "encoded protein shown in AAW72185"
 FT complement (26295..47251)
 FT /*tag= z
 FT /product= "ORF#19 protein"
 FT /note= "encoded protein shown in AAW72186"
 FT complement (27630..31754)
 FT /*tag= aa
 FT /product= "ORF#20b protein"
 FT /note= "encoded protein shown in AAW72225"
 FT complement (27630..31784)
 FT /*tag= ab
 FT /product= "ORF#20a protein"
 FT /note= "encoded protein shown in AAW72224"
 FT complement (32067..32735)
 FT /*tag= ac
 FT /product= "ORF#21 protein"
 FT /note= "encoded protein shown in AAW72187"
 FT 33140..34984
 FT /*tag= ad
 FT /product= "ORF#22a protein"
 FT /note= "encoded protein shown in AAW72226"
 FT 33386..34984
 FT /*tag= ae
 FT /product= "ORF#22b protein"
 FT /note= "encoded protein shown in AAW72227"
 FT complement (35205..37721)
 FT /*tag= af
 FT /product= "ORF#23 protein"
 FT /note= "encoded protein shown in AAW72188"
 FT complement (38058..39188)
 FT /*tag= ag
 FT /product= "ORF#24 protein"
 FT /note= "encoded protein shown in AAW72189"
 FT 39090..39935
 FT /*tag= ah
 FT /product= "ORF#25 protein"
 FT /note= "encoded protein shown in AAW72190"
 FT 40216..41973
 FT /*tag= ai
 FT /product= "ORF#26 protein"
 FT /note= "encoded protein shown in AAW72191"
 FT 42206..44178
 FT /*tag= aj
 FT /transl_except= (pos: 44063..44064, aa: Lys)
 FT /product= "ORF#27 protein"
 FT /note= "encoded protein shown in AAW72192"
 FT complement (44853..47297)
 FT /*tag= ak
 FT /product= "ORF#28 protein"
 FT /note= "encoded protein shown in AAW72193"
 FT 47122..47338

```
FT      /tag= al
FT      /product= "ORF#29 protein"
FT      /note= "encoded protein shown in AAW72194"
FT      complement (47305..49662)
FT      /tag= am
FT      /product= "ORF#30 protein"
FT      /note= "encoded protein shown in AAW72195"
FT      complement (51035..51666)
FT      /tag= an
FT      /product= "ORF#31 protein"
FT      /note= "encoded protein shown in AAW72196"
FT      complement (51701..53575)
FT      /tag= ao
FT      /product= "ORF#32 protein"
FT      /note= "encoded protein shown in AAW72197"
FT      complement (54393..58115)
FT      /tag= ap
FT      /product= "ORF#33 protein"
FT      /note= "encoded protein shown in AAW72198"
FT      complement (58060..58977)
FT      /tag= aq
FT      /product= "ORF#34 protein"
FT      /note= "encoded protein shown in AAW72199"
FT      complement (58970..60760)
FT      /tag= ar
FT      /product= "ORF#35 protein"
FT      /note= "encoded protein shown in AAW72200"
FT      complement (60759..61151)
FT      /tag= as
FT      /product= "ORF#36 protein"
FT      /note= "encoded protein shown in AAW72201"
FT      complement (61241..62071)
FT      /tag= at

Query Match
Best Local Similarity 68.5%; Score 22.6; DB 19; Length 117213;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      6 CAAACTCGTGTCTCTCCAGCATG 28
DB      45257 CAAACTCGTGTCTCTCCAGCATG 45279

RESULT 9
AAD25519
ID      AAD25519 standard; DNA; 154746 BP.
XX      AAD25519;
AC      AAD25519;
XX      AAD25519;
DT      26-MAR-2002 (first entry)
XX      26-MAR-2002 (first entry)
DE      Human herpesvirus 2 complete DNA genome.
XX      Human herpesvirus 2; cytostatic; cancer; immunosuppressive; virucide;
XX      antibacterial; fungicide; protozoicide; antirheumatic; antiinflammatory;
XX      antitachytic; rheumatoid arthritis; neuroprotective; multiple sclerosis;
XX      immune response; vasectropic; vaccine; gene therapy; autoimmune disease;
XX      vasculitis; de.
XX      Human herpesvirus 2.
OS      Human herpesvirus 2.
XX      WO200176643-A1.
XX      WO200176643-A1.
PD      18-OCT-2001.
XX      18-OCT-2001.
PF      06-APR-2001; 2001WO-US11372.
XX      06-APR-2001; 2001WO-US11372.
XX      07-APR-2000; 2000US-195680P.
XX      07-APR-2000; 2000US-195680P.
PA      (BAYU ) BAYLOR COLLEGE MEDICINE.
XX      (BAYU ) BAYLOR COLLEGE MEDICINE.
PI      Orson FM, Kinney BM, Bhogal BS;
XX      Orson FM, Kinney BM, Bhogal BS;
```

```
DR      WPI; 2002-066308/09.
XX      Composition for oral delivery of vaccines, comprises expression vector
XX      containing antigenic genomic sequence, bound to aggregated
XX      protein-polycationic polymer conjugate or suspension
XX      protein-polycationic polymer conjugate or suspension
PS      Disclosure; Page 90-132; 145pp; English.
XX      The invention relates to a composition comprising an expression vector
XX      bound to an aggregated protein-polycationic polymer conjugate or
XX      suspension. The expression vector contains a promoter polynucleotide
XX      sequence operatively linked to a polynucleotide sequence encoding an
XX      antigen which is a fragment of a gene or genome associated with an
XX      infectious disease, cancer and autoimmune disease such as rheumatoid
XX      arthritis, vasculitis, and multiple sclerosis, pathogenic genomes
XX      consisting of bacterium, fungus, protozoa and virus such as human
XX      immunodeficiency virus (HIV), herpes simplex virus (HSV), hepatitis C
XX      virus (HCV), influenza and respiratory syncytial virus (RSV), and
XX      optionally comprising a nucleotide sequence encoding a cytokine (or a
XX      cytokine expression vector), is useful for inducing an immune response
XX      (systemic and/or mucosal) in an organism. The cytokine expression vector
XX      contains a sequence for granulocyte macrophage-colony stimulating factor
XX      (GM-CSF) or interleukin-12 (IL-12). The polynucleotide sequences encoding
XX      the antigen and the cytokine are under transcriptional control of same or
XX      different promoter polynucleotide sequences. The expression vector, as a
XX      CC DNA vaccine is useful for treating a condition in an organism. The
XX      CC present sequence is human herpesvirus 2 complete DNA genome related
XX      to the invention.
SO      Sequence 154746 BP; 23003 A; 54218 C; 54701 G; 22824 T; 0 other;

Query Match
Best Local Similarity 68.5%; Score 22.6; DB 24; Length 154746;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      6 CAAACTCGTGTCTCTCCAGCATG 28
DB      54077 CAAACTCGTGTCTCTCCAGCATG 54099

RESULT 10
AA161247/c
ID      AA161247 standard; cDNA; 2550 BP.
XX      AA161247;
AC      AA161247;
XX      AA161247;
DT      22-OCT-2001 (first entry)
XX      22-OCT-2001 (first entry)
DE      Human polynucleotide SEQ ID NO 5236.
XX      Human polynucleotide SEQ ID NO 5236.
XX      Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX      peripheral nervous system; neuropathy; central nervous system; CNS;
XX      Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX      amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX      chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX      leukaemia; ss.
XX      Homo sapiens.
OS      Homo sapiens.
XX      WO200153312-A1.
XX      WO200153312-A1.
PD      26-JUL-2001.
XX      26-JUL-2001.
PF      26-DEC-2000; 2000WO-US34263.
XX      26-DEC-2000; 2000WO-US34263.
XX      21-JAN-2000; 2000US-0488725.
XX      21-JAN-2000; 2000US-0488725.
XX      25-APR-2000; 2000US-0552317.
XX      25-APR-2000; 2000US-0552317.
XX      09-JUL-2000; 2000US-0598042.
XX      09-JUL-2000; 2000US-0598042.
XX      19-JUL-2000; 2000US-0620312.
XX      19-JUL-2000; 2000US-0620312.
XX      03-AUG-2000; 2000US-0653450.
XX      03-AUG-2000; 2000US-0653450.
XX      14-SEP-2000; 2000US-0662191.
XX      14-SEP-2000; 2000US-0662191.
XX      19-OCT-2000; 2000US-0693036.
XX      19-OCT-2000; 2000US-0693036.
XX      29-NOV-2000; 2000US-0727344.
XX      29-NOV-2000; 2000US-0727344.
```

XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Auendi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR P-PSDB; AAM42091.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS Claim 1; SEQ ID NO 5236; 10078bp; English.
 XX The invention relates to human nucleic acids (AA15798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nucleotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: Immune system suppression,
 CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 XX Sequence 2550 BP; 633 A; 686 C; 650 G; 581 T; 0 other;
 SQ
 XX Query Match 67.9%; Score 22.4; DB 22; Length 2550;
 Best Local Similarity 88.5%; Pred. No. 21;
 Matches 23; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 4 TGCAAACTCGTGTCTCCAGCATGC 29
 Db 754 TGCAAACTCGTGTCTCCAGCATGC 729

RESULT 11
 AA153531/c
 ID AA153531 standard; cDNA; 2815 BP.
 XX
 AC AA153531;
 XX
 DT 30-JAN-2003 (first entry)
 XX
 DE cDNA of human EHD2 gene, SEQ ID NO 7.
 XX
 KW Cytostatic; osteopathic; recombinant Eps15 homology; EH domain; EHD1;
 KW EHD2; endocytosis; IGf1 signaling; suppressing adhesion; invasion;
 KW metastasis; bone formation; osteoporosis; human; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 51..1001
 FT /*tag= a
 FT /product= "Human EHD2 protein"
 XX
 PN US2002115069-A1.
 XX
 XX 22-AUG-2002.
 PD
 XX 17-MAY-1999; 99US-0312762.
 PF
 XX 21-FEB-1997; 97IL-0120283.
 PR
 XX 20-FEB-1998; 98US-0026898.
 PR
 XX

PA (HORO/) HOROWITZ M.
 PA (MINT/) MINTZ L.
 XX
 PI Horowitz M, Mintz L;
 XX
 DR WPI; 2002-731288/79.
 DR
 DR P-PSDB; AAO26366.
 XX
 XX Novel recombinant Eps15 homology domain containing protein useful for
 PT regulating endogenous EHD protein by regulating insulin-like growth
 PT factor 1 receptor cell signaling via altered clathrin coated pit
 PT mediated endocytosis -
 PS Claim 6; Page 45-46; 88bp; English.
 XX
 XX The invention relates to a novel recombinant Eps15 homology (EH) domain
 CC containing protein 1 or 2 (EHD1 or EHD2) with an N-terminal region
 CC containing a nucleotide binding consensus site, a central coiled coil
 CC structure and a C-terminal region including an EH domain, where the
 CC polypeptide participates in endocytosis. The peptides, peptide analogues
 CC and/or the polynucleotide sequences are useful for regulating EHD1/2
 CC protein activity in vivo, where the EHD1/2 protein and its gene are
 CC useful for upregulating the protein activity. The novel method is useful
 CC for lowering the rate of IGf1 signaling and therefore useful for
 CC suppressing adhesion, invasion and metastasis where EHD is overexpressed,
 CC and therefore endocytosis. The method is also useful for elongating IGf1
 CC effects and therefore for increasing bone formation in osteoporosis
 CC conditions. This polynucleotide sequence represents the cDNA of a human
 CC EHD2 gene of the invention.
 XX
 XX Sequence 2815 BP; 695 A; 754 C; 730 G; 636 T; 0 other;
 SQ
 XX Query Match 67.9%; Score 22.4; DB 24; Length 2815;
 Best Local Similarity 88.5%; Pred. No. 21;
 Matches 23; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 4 TGCAAACTCGTGTCTCCAGCATGC 29
 Db 980 TGCAAACTCGTGTCTCCAGCATGC 955

RESULT 12
 AA159461/c
 ID AA159461 standard; cDNA; 3327 BP.
 XX
 AC AA159461;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 1664.
 XX
 KW Human; nucleotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 51..1001
 FT /*tag= a
 FT /product= "Human EHD2 protein"
 XX
 PN WO200153312-A1.
 XX
 XX 26-JUL-2001.
 PD
 XX 26-DEC-2000; 2000WO-US34263.
 PF
 XX 21-JAN-2000; 2000US-0488725.
 PR
 XX 25-APR-2000; 2000US-0552317.
 PR
 XX 09-JUL-2000; 2000US-0588042.
 PR
 XX 19-JUL-2000; 2000US-0620312.
 PR
 XX 03-AUG-2000; 2000US-0653450.
 PR
 XX 14-SEP-2000; 2000US-0662191.
 PR
 XX 19-OCT-2000; 2000US-0693036.
 PR

```

XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Aundt V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
XX Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Demanac RT;
XX
XX WPI: 2001-443253/47.
XX P-PSDB; AAM40305.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX PT such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 1664; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX CC the encoded polypeptides (AAM38642-AAW42213) with nootropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: immune system suppression,
XX CC activation/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and
XX CC C.N.S disorders.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification.
XX
XX Sequence 3327 BP; 814 A; 900 C; 880 G; 733 T; 0 other;
XX
XX
XX Query March 67.9%; Score 22.4; DB 22; Length 3327;
XX Best Local Similarity 88.5%; Pred. No. 22;
XX Matches 23; Conservative 1; Mismatches 2; Indels 0; Gaps 0
XX
XX 4 TGCAAACTCGTGTCTCTCCAGCATGC 29
XX Db 1522 TGCAAACTCGTGTCTCTCCAGCATGC 1497
XX
XX
XX RESULT 13
XX ABZ41610/C
XX ID ABZ41610 standard; DNA; 1494 BP.
XX
XX ABZ41610;
XX
XX 07-MAR-2003 (first entry)
XX
XX N. gonorrhoeae nucleotide sequence SEQ ID 7809.
XX
XX Antibacterial, infection; vaccine; gene therapy; gene; ds.
XX
XX Neisseria gonorrhoeae.
XX
XX WO200279243-A2.
XX
XX 10-OCT-2002.
XX
XX 12-FEB-2002; 2002WO-IB02069.
XX
XX 12-FEB-2001; 2001GB-0003424.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Fontana MR, Pizsa M, Maignani V, Monaci E;
XX WPI: 2003-058415/05.
XX P-PSDB; ABP80640.
XX

```

PT	New protein from Neisseria gonorrhoeae, useful for the manufacture of a medicament for treating or preventing N. gonorrhoeae infection -
PS	Disclosure; Page 758; 815pp; English.
XX	The present invention relates to proteins from Neisseria gonorrhoeae. CC
CC	Also disclosed are the nucleic acid molecules encoding the proteins and CC
CC	antibodies that specifically bind to the proteins. The composition CC
CC	comprising the protein, nucleic acid or antibody is useful for the CC
CC	manufacture of a medicament for treating or preventing N. gonorrhoeae CC
CC	infection. This may be in the form of a vaccine or gene therapy. CC
CC	Sequences given in records ABX37706-ABX42016 represent nucleic acid CC
XX	molecules of the invention.
SQ	Sequence 1494 BP; 304 A; 453 C; 432 G; 305 T; 0 other;
OY	Query Match 64.8%; Score 21.4; DB 25; Length 1494; Best Local Similarity 75.8%; Pred. No. 50; Matches 25; Conservative 1; Mismatches 7; Indels 0; Gaps 0.
DB	1 CCCGCAACCTCGTGTCTTCAGCATGCAGGG 33 324 CCCGCCAACCTCGATTTCCGCCACTTTCCGG 292
RESULT 14	
ID	AAO14479/C
XX	AAO14479 standard; DNA; 2088 BP.
AC	AAO14479;
XX	
DT	25-MAR-2003 (updated)
DT	23-JUN-1992 (first entry)
XX	
DE	Truncated HSVB gene.
XX	
KW	Vaccine; antigen; ss.
OS	Herpes simplex virus.
FT	Key Location/Qualifiers
FT	CDS 1..2088
FT	/tag= a
PN	JF03218397-A.
XX	
PD	25-SEP-1991.
XX	
PE	21-JUN-1990; 90JP-0161448;
XX	
PR	30-NOV-1989; 89JP-0308941.
PR	22-JUN-1989; 89JP-0158238.
PR	21-JUN-1990; 90JP-0161448.
XX	
PA	(TAKEDA) TAKEDA CHEM IND LTD.
DR	WI; 1991-328397/45.
DR	P-PADB; AARI4666.
XX	
FT	HSVGB polypeptide(s) obtd. by recombinant DNA techniques -
PT	useful as vaccines against HSV and in diagnosis, can be produced
PT	cheaply and safely.
XX	
PS	Claim 3; Fig 7; 24pp; Japanese.
CC	The sequence encodes a truncated form of the HSVGB polypeptide.
CC	The recombinant protein can be used to prepare vaccines for
CC	prophylaxis of HSV infection and for use in diagnostic kits.
CC	See also AAO14478.
CC	(Updated on 25-MAR-2003 to correct PA field.)
SQ	Sequence 2088 BP; 417 A; 751 C; 620 G; 300 T; 0 other;

Query Match 64.8%; Score 21.4; DB 12; Length 2088;
 Best Local Similarity 95.7%; Pred. No. 52;
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 CAAACTCGTGTCTCTCCGAGCATG 28
 |||||
 DB 1963 CAAACTCGTGTCTCTCCGAGCATG 1941

RESULT 15

AA71303/C
 ID AA71303 standard; DNA; 2712 BP.

XX AA71303;

AC 25-MAR-2003 (updated)

DT 30-APR-1991 (first entry)

XX Herpes Simplex Virus-1 GB from PRBXX.

DE Vaccine; prophylaxis; treatment; Herpes Simplex Virus-1;

XX glycoprotein; gb; ss.

KM Herpes simplex virus type 1 (KOS).

XX US464233-A.

XX 10-FEB-1987.

PD 20-JUN-1984; 84US-0622496.

XX 20-JUN-1984; 84US-0622496.

XX (PERS/) PERSON S.

PA Person S;

PT WPI; 1987-056354/08.

XX P-PSDB; AAP7135.

DR Amino acid chain of glycoprotein B of HSV-1 and 2 - prepd. as

XX recombinant and used for vaccines for herpes simplex virus types 1

PT and 2.

XX Example; Table 2; 16pp; English.

XX A pure non-glycosylated amino acid (AA) chain comprising a sequence

CC corresponding to that occurring in glycoprotein B of HSV-1 or HSV-2

CC virus which is antigenic to HSV-1 of HSV-2, which contains no more

CC than 750 AA residues, and which includes AA residues 135-649

CC inclusive is claimed. It can be used to produce vaccines for

CC prophylaxis and treatment of HSV-1 and HSV-2.

CC See AA71399 for the HSV-2 sequence.

CC (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 2712 BP; 519 A; 945 C; 849 G; 397 T; 2 other;

XX Query Match 64.8%; Score 21.4; DB 8; Length 2712;

XX Best Local Similarity 95.7%; Pred. No. 54;

XX Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 CAAACTCGTGTCTCTCCGAGCATG 28

|||||

DB 2047 CAAACTCGTGTCTCTCCGAGCATG 2025

Search completed: September 16, 2003, 18:37:07
 Job time : 167.327 secs

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 18:24:35 ; Search time 42.1485 Seconds
(without alignments)
345.579 Million cell updates/sec

Title: US-09-594-065-4
Perfect score: 33
Sequence: 1 ccctgcaactcgtgkctccacgacatgcagg 33

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB	ID	Description
C 1	22.6	68.5	3472	6	Patent No. 5244792-1
C 2	22.4	67.9	2815	4	US-09-312-762A-7
C 3	21.4	64.8	2713	3	US-08-804-439A-13
C 4	21.4	64.8	2713	3	US-08-720-229-13
C 5	21.4	64.8	3472	6	5244792-2
C 6	20.2	61.2	2517	1	US-07-906-930E-1
C 7	20.2	61.2	2939	1	US-07-906-930E-3
C 8	20.2	61.2	3138	4	US-09-434-408-1
C 9	19.4	58.8	4403765	3	US-09-103-840A-2
C 10	19.4	58.8	4411529	3	US-09-103-840A-1
C 11	19.2	58.2	1001	4	US-09-641-638-358
C 12	19.2	58.2	1001	4	US-09-641-638-359
C 13	18.8	57.0	1656	4	US-09-904-615-27
C 14	18.8	57.0	3955	3	US-09-214-278-4
C 15	18.8	57.0	4464	2	US-08-400-159-7
C 16	18.8	57.0	4483	3	US-08-611-729A-7
C 17	18.6	56.4	1176	4	US-09-663-600A-42
C 18	18.6	56.4	1887	4	US-09-149-727-9
C 19	18.6	56.4	2573	4	US-09-620-312D-609
C 20	18.6	56.4	1230025	4	US-09-198-452A-1
C 21	18.4	55.8	2943	1	US-08-042-747A-7
C 22	18.4	55.8	3177	1	US-08-042-747A-4
C 23	18.2	55.2	3348	4	US-09-312-762A-2
C 24	18.2	55.2	3635	4	US-09-312-762A-6
C 25	18.2	55.2	7812	3	US-09-368-590-1
C 26	18.2	55.2	8453	3	US-09-167-681-45
C 27	18.2	55.2	14707	4	US-09-312-762A-3

28	18	54.5	1112 4 US-09-663-600A-136	Sequence 136, App
29	17.8	53.9	2311 2 US-08-976-259-26	Sequence 26, Appli
30	17.8	53.9	3627 2 US-08-232-087A-1	Sequence 1, Appli
31	17.8	53.9	5084 1 US-08-306-691B-21	Sequence 21, Appli
32	17.8	53.9	5084 5 PCT-US93-06251-25	Sequence 25, Appli
33	17.6	53.3	624 4 US-09-221-017B-43	Sequence 43, Appli
34	17.6	53.3	795 4 US-09-252-991A-3803	Sequence 3803, Ap
35	17.6	53.3	1122 3 US-08-746-883-1	Sequence 1, Appli
36	17.6	53.3	1623 4 US-09-252-991A-3844	Sequence 3844, Ap
37	17.6	53.3	4167 4 US-09-252-991A-3666	Sequence 3666, Ap
38	17.6	53.3	8802 3 US-08-896-449A-1	Sequence 1, Appli
39	17.6	53.3	8802 3 US-09-132-652-1	Sequence 1, Appli
40	17.4	52.7	576 4 US-09-252-991A-13853	Sequence 13853, A
41	17.4	52.7	1098 4 US-09-252-991A-13595	Sequence 13595, A
42	17.4	52.7	1340 4 US-09-000-062-2	Sequence 2, Appli
43	17.4	52.7	1340 4 US-09-000-062-2	Sequence 4, Appli
44	17.4	52.7	1340 4 US-08-945-144A-2	Sequence 2, Appli
45	17.4	52.7	1340 4 US-08-945-144A-4	Sequence 4, Appli

ALIGNMENTS

```

RESULT 1
5244792-1/c
Patent No. 5244792
APPLICANT: BURKE, RAE L.; PACHL, CAROL; VALENZUELA, PABLO D.T.
TITLE OF INVENTION: EXPRESSION OF RECOMBINANT GLYOPIROTEIN
B FROM HERPES SIMPLEX VIRUS
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587,179
FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 921,730
FILING DATE: 20-OCT-1986
APPLICATION NUMBER: 597,784
FILING DATE: 06-APR-1984
SEQ ID NO:1:
LENGTH: 3472
5244792-1

Query Match      68.5%; Score 22.6; DB 6; Length 3472;
Best Local Similarity 95.7%; Pred. No. 1.6;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy      6 CAAACTGCTGCTCTCCAGCATG 28
Db      2349 CAAACTGCTGCTCTCCAGCATG 2327

RESULT 2
US-09-312-762A-7/c
Sequence 7, Application US/09312762A
Patent No. 6552177
GENERAL INFORMATION:
APPLICANT: MIA HOROWITZ ET AL.
TITLE OF INVENTION: EH DOMAIN CONTAINING GENES AND PROTEINS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSER: Mark M. Friedman c/o Anthony Caetorina
STREET: 2001 Jefferson Davis Highway, Suite 207
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
COMPUTER: Twinhead* Slimnote-890TX
OPERATING SYSTEM: MS DOS version 6.2,
OPERATING SYSTEM: Windows version 3.11
SOFTWARE: Word for Windows version 2.0 converted to
SOFTWARE: an ASCII file

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/312,762A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/026,898
FILING DATE: 20 FEB 1998
ATTORNEY/AGENT INFORMATION:
NAME: Friedman, Mark M.
REGISTRATION NUMBER: 33,883
REFERENCE/DOCKET NUMBER: 916/10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 972-3-5625553
TELEFAX: 972-3-5625554
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2815
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-312-762A-7

Query Match 67.9%; Score 22.4; DB 4; Length 2815;
Best Local Similarity 88.5%; Pred. No. 1.9;
Matches 23; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 TGCACCTGCTGCTCTCCAGCATG 29
DB 980 TGCACCTGCTGCTCTCCAGCATG 955

RESULT 3
US-08-804-439A-13/c
Sequence 13, Application US/08804439A
Patent No. 6015565
GENERAL INFORMATION:
APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Mark L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHV/KSHV
TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Ste 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,439A
FILING DATE: February 21, 1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Hall, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09176/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
TELEX:
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2713 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

US-08-804-439A-13
Query Match 64.8%; Score 21.4; DB 3; Length 2713;
Best Local Similarity 95.7%; Pred. No. 5;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 CAACCTGCTGCTCTCCAGCATG 28
DB 2047 CAACCTGCTGCTCTCCAGCATG 2025

RESULT 4
US-08-720-229-13/c
Sequence 13, Application US/08720229
Patent No. 6022542
GENERAL INFORMATION:
APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Mark L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHV/KSHV
TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,229
FILING DATE: 26-SEP-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2713 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-720-229-13

Query Match 64.8%; Score 21.4; DB 3; Length 2713;
Best Local Similarity 95.7%; Pred. No. 5;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 CAACCTGCTGCTCTCCAGCATG 28
DB 2047 CAACCTGCTGCTCTCCAGCATG 2025

RESULT 5
5244792-2/c
Patent No. 5244792
APPLICANT: BURKE, RAE L.; PACHL, CAROL; VALENZUELA, PABLO D.T.
TITLE OF INVENTION: EXPRESSION OF RECOMBINANT GLYCOPROTEIN
B FROM HERPES SIMPLEX VIRUS
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587,179
FILING DATE: 20-SEP-1990

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 921,730
FILING DATE: 20-OCT-1986
APPLICATION NUMBER: 597,784
FILING DATE: 06-APR-1984
SEQ ID NO: 2:
LENGTH: 3472
5244792-2

Query Match 64.8%; Score 21.4; DB 6; Length 3472;
Best Local Similarity 95.7%; Pred. No. 5.2;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 CAAACTGCTGTCCTCCAGCATG 28
DB 2358 CAAACTGCTGTCCTCCAGCATG 2336

RESULT 6
US-07-906-930E-1
Sequence 1, Application US/07906930E
Patent No. 5534631
GENERAL INFORMATION:
APPLICANT: Gaynor, Richard B.
APPLICANT: Nirula, Ajay
TITLE OF INVENTION: DNA ENCODING THE INTERLEUKIN BINDING
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/906,930E
FILING DATE: 30-JUN-1992
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Serlich, Gary J.
REGISTRATION NUMBER: 34,430
REFERENCE/DOCKET NUMBER: UTSD:262/SER
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-418-3000
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2517 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
US-07-906-930E-1

Query Match 61.2%; Score 20.2; DB 1; Length 2517;
Best Local Similarity 75.8%; Pred. No. 16;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCCTGCAAACTCGTGTCTCCAGCATG 33
DB 255 CGCTGCAAACTCGTGTCTCCAGCATG 287

RESULT 7

US-07-906-930E-3

Sequence 3, Application US/07906930E
Patent No. 5534631
GENERAL INFORMATION:
APPLICANT: Gaynor, Richard B.
APPLICANT: Nirula, Ajay
TITLE OF INVENTION: DNA ENCODING THE INTERLEUKIN BINDING
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/906,930E
FILING DATE: 30-JUN-1992
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Serlich, Gary J.
REGISTRATION NUMBER: 34,430
REFERENCE/DOCKET NUMBER: UTSD:262/SER
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-418-3000
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2939 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
US-07-906-930E-3

Query Match 61.2%; Score 20.2; DB 1; Length 2939;
Best Local Similarity 75.8%; Pred. No. 16;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCCTGCAAACTCGTGTCTCCAGCATG 33
DB 255 CGCTGCAAACTCGTGTCTCCAGCATG 287

RESULT 8
US-09-434-408-1/C
Sequence 1, Application US/09434408
Patent No. 6440697
GENERAL INFORMATION:
APPLICANT: Venezia, Domenick
APPLICANT: Grossmann, Angelika
TITLE OF INVENTION: RING FINGER PROTEIN ZAPOR3
FILE REFERENCE: 98-41
CURRENT APPLICATION NUMBER: US/09/434,408
CURRENT FILING DATE: 1999-11-04
EARLIER APPLICATION NUMBER: US 60/108,258
EARLIER FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 3138
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: CDS
LOCATION: (367)...(2535)
US-09-434-408-1

Query Match 61.2%; Score 20.2; DB 4; Length 3138;
Best Local Similarity 81.5%; Pred. No. 16;
Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCTGCAAACTCGTGTCTCCAGCAT 27
DB 1162 CCTGCAAACTCGAGTTCTCCAGCAT 1156

RESULT 9
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6284328

GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.

APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103,840A

CURRENT FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 4403765

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis

FEATURE:

OTHER INFORMATION: CDC 1551

OTHER INFORMATION: "n" bases at various positions throughout the sequence

OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Query Match 58.8%; Score 19.4; DB 3; Length 4403765;
Best Local Similarity 74.2%; Pred. No. 72;
Matches 23; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 CTGCAAACTCGTGTCTCCAGCATGCAAGG 33
DB 2096169 CGGCAACGCGAGGTCTCTCCGAGTCCGGCG 2096139

RESULT 10
US-09-103-840A-1/c
Sequence 1, Application US/09103840A
Patent No. 6284328

GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.

APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103,840A

CURRENT FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 4411529

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis

OTHER INFORMATION: H37Rv

US-09-103-840A-1

Query Match 58.8%; Score 19.4; DB 3; Length 4411529;
Best Local Similarity 74.2%; Pred. No. 72;

Matches 23; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 3 CTGCAAACTCGTGTCTCCAGCATGCAAGG 33
DB 2098979 CGGCAACGCGAGGTCTCTCCGAGTCCGGCG 2098949

RESULT 11
US-09-641-638-358
Sequence 358, Application US/09641638
Patent No. 6432648

GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta

APPLICANT: Bougueleret, Lydie

APPLICANT: Chumakov, Ilya

APPLICANT: Cohen, Annick

TITLE OF INVENTION: BIALLIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING

FILE REFERENCE: GENSET.051CP1

CURRENT APPLICATION NUMBER: US/09/641,638

CURRENT FILING DATE: 2000-08-16

PRIOR APPLICATION NUMBER: US 09/502,330

PRIOR FILING DATE: 2000-02-11

PRIOR APPLICATION NUMBER: US 60/133,200

PRIOR FILING DATE: 1999-05-07

PRIOR APPLICATION NUMBER: US 09/275,267

PRIOR FILING DATE: 1999-03-23

PRIOR APPLICATION NUMBER: US 60/119,917

PRIOR FILING DATE: 1999-02-12

NUMBER OF SEQ ID NOS: 1304

SOFTWARE: Patent.pm

SEQ ID NO 358

LENGTH: 1001

TYPE: DNA

ORGANISM: Homo Sapiens

FEATURE:

NAME/KEY: allele

LOCATION: 501

OTHER INFORMATION: 12-866-423 : polymorphic base C or T

NAME/KEY: misc_binding

LOCATION: 482..500

OTHER INFORMATION: 12-866-423.mis1

NAME/KEY: misc_binding

LOCATION: 502..521

OTHER INFORMATION: 12-866-423.mis2, potential complement

NAME/KEY: primer_bind

LOCATION: 79..98

OTHER INFORMATION: upstream amplification primer

NAME/KEY: primer_bind

LOCATION: 589..609

OTHER INFORMATION: downstream amplification primer, complement

NAME/KEY: misc_binding

LOCATION: 489..513

OTHER INFORMATION: 12-866-423 potential probe

US-09-641-638-358

Query Match 58.2%; Score 19.2; DB 4; Length 1001;
Best Local Similarity 80.8%; Pred. No. 36;
Matches 21; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 CTGCAAACTCGTGTCTCCAGCATG 28
DB 906 CTGCAAACTCGTGTCTCCAGCATG 931

RESULT 12
US-09-641-638-359

Sequence 359, Application US/09641638
Patent No. 6432648

GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta

APPLICANT: Bougueleret, Lydie

APPLICANT: Chumakov, Ilya

APPLICANT: Cohen, Annick

TITLE OF INVENTION: BIALLIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING

FILE REFERENCE: GENSET.051CP1

CURRENT APPLICATION NUMBER: US/09/641,638

CURRENT FILING DATE: 2000-08-16

PRIOR APPLICATION NUMBER: US 09/502,330

PRIOR FILING DATE: 2000-02-11

PRIOR APPLICATION NUMBER: US 60/133,200

PRIOR FILING DATE: 1999-05-07

PRIOR APPLICATION NUMBER: US 09/275,267

PRIOR FILING DATE: 1999-03-23

PRIOR APPLICATION NUMBER: US 60/119,917

PRIOR FILING DATE: 1999-02-12

NUMBER OF SEQ ID NOS: 1304

SOFTWARE: Patent.pm

SEQ ID NO 358

LENGTH: 1001

TYPE: DNA

ORGANISM: Homo Sapiens

FEATURE:

NAME/KEY: allele

LOCATION: 501

OTHER INFORMATION: 12-866-423 : polymorphic base C or T

NAME/KEY: misc_binding

LOCATION: 482..500

OTHER INFORMATION: 12-866-423.mis1

NAME/KEY: misc_binding

LOCATION: 502..521

OTHER INFORMATION: 12-866-423.mis2, potential complement

NAME/KEY: primer_bind

LOCATION: 79..98

OTHER INFORMATION: upstream amplification primer

NAME/KEY: primer_bind

LOCATION: 589..609

OTHER INFORMATION: downstream amplification primer, complement

NAME/KEY: misc_binding

LOCATION: 489..513

OTHER INFORMATION: 12-866-423 potential probe

US-09-641-638-358

```

1  APPLICANT: Cohen, Annick
2  TITLE OF INVENTION: BIALLIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
3  TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
4  FILE REFERENCE: GENSET.051CPI
5  CURRENT APPLICATION NUMBER: US/09/641,638
6  PRIOR FILING DATE: 2000-08-16
7  PRIOR APPLICATION NUMBER: US 09/502,330
8  PRIOR FILING DATE: 2000-02-11
9  PRIOR APPLICATION NUMBER: US 60/133,200
10 PRIOR FILING DATE: 1999-05-07
11 PRIOR APPLICATION NUMBER: US 09/275,267
12 PRIOR FILING DATE: 1999-03-23
13 PRIOR APPLICATION NUMBER: US 60/119,917
14 PRIOR FILING DATE: 1999-02-12
15 NUMBER OF SEQ ID NOS: 1304
16 SOFTWARE: Patent.pm
17 SEQ ID NO 359
18 LENGTH: 1001
19 TYPE: DNA
20 ORGANISM: Homo Sapiens
21 FEATURE:
22 NAME/KEY: allele
23 LOCATION: 501
24 OTHER INFORMATION: 12-867-47 : polymorphic base C or T
25 NAME/KEY: misc_binding
26 LOCATION: 481..500
27 OTHER INFORMATION: 12-867-47.misl, potential
28 NAME/KEY: misc_binding
29 LOCATION: 502..521
30 OTHER INFORMATION: 12-867-47.mis2, potential complement
31 NAME/KEY: primer_bind
32 LOCATION: 455..474
33 OTHER INFORMATION: upstream amplification primer
34 NAME/KEY: primer_bind
35 LOCATION: 898..918
36 OTHER INFORMATION: downstream amplification primer, complement
37 NAME/KEY: misc_binding
38 LOCATION: 489..513
39 OTHER INFORMATION: 12-867-47 potential probe
40 US-09-641-638-359

```

Query Match	58.2%	Score 19.2	DB 4	Length 1001
Best Local Similarity	80.8%	Pred. No. 36		
Matches	21	Conservative	1	Mismatches 4; Indels 0; Gaps 0;
Qy	3	CTGCAAACTCGTGCTCTCCAGCATG	28	
Db	466	CTGCAAACTCTGTCTCCAGCATG	491	

```

RESULT 13
US-09-904-615-27/c
; Sequence 27. Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032p1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 1656
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-904-615-27

```

Query Match	57.0%	Score 18.8;	DB 4;	Length 1656;
Best Local Similarity	71.9%	Pred. No. 58;		
Matches 23;	Conservative 1;	Mismatches 8;	Indels 0;	Gaps 0;

2 CCGCAACCTGCTGTCCTCCAGCATGAGGG 33
 231 CGTGCMACTCCACGGGTTCCAGCATGCAATG 200

```

RESULT 14
US-09-214-278-4/c
; Sequence 4, Application US/09214278
; Patent No. 6291210
; GENERAL INFORMATION:
; APPLICANT: Sakano, Sei-I
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-0576
; CURRENT APPLICATION NUMBER: US/09/214,278
; CURRENT FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3955
; MAKE DNA

```

```

? ORGANISM: Homo sapiens
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (12)..(3725)
?
? NAME/KEY: sig_peptide
? LOCATION: (12)..(89)
?
? NAME/KEY: mat_peptide
? LOCATION: (90)..(3725)
?
? US-09-214-278-4

```

Query Match	57.0%;	Score 18.8;	DB 3;	length 3955;
Best Local Similarity	71.9%;	Pred. No. 67;		
Matches	23;	Conservative	1;	Mismatches 8;
				Indels 0;
				Gaps 0;

Oy		1	CCCTGCAAACTCGTGGKTCCTCCAGCATGTAGG	32
			:	
Ddb		1288	CCCTTCACACTCATTGGCGTCCAGCTGGCAGG	1267

US-08T 15
Sequence 7, Application US/08400159
Patent No. 5869282
GENERAL INFORMATION:
APPLICANT: Ish-Horowitz, David
APPLICANT: Henriquez, Domingos M.P.
APPLICANT: Lewis, Julian H.
APPLICANT: Myat, Anna M.
APPLICANT: Fleming, Robert J.
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Mann, Robert S.
APPLICANT: Gray, Grace E.
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
TYPE OF INVENTION: SRRATE GENE AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/400,159
/ FILING DATE: 07-MAR-1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mirock, S. Leslie
/ REGISTRATION NUMBER: 18,872
/ REFERENCE/DOCKET NUMBER: 7326-029
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-9741/8864
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 464 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: unknown
/ MOLECULE TYPE: cDNA
/ US-08-400-159-7

Query Match 57.0%; Score 18.8; DB 2; Length 464;
Best Local Similarity 71.9%; Pred. No. 68;
Matches 23; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CCTGCAACTGCTGTCCTCCGACATGCAGG 32
||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1108 CCTTCACACTCATTTGGCGCTCCAGCTGCAGG 1077

Search completed: September 16, 2003, 20:42:50
Job time : 54.1485 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 19:14:25 ; Search time 119.257 Seconds
(without alignments)
680.690 Million cell updates/sec

Title: US-09-594-065-4

Perfect score: 33

Sequence: 1 cccctgcaactcgtgkctccagcagcaggg 33

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1660708 seqs, 1229959015 residues

Total number of hits satisfying chosen parameters: 3321416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US00_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US00_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	22.6	68.5	154746	12	US-09-827-688-8
2	22.4	67.9	639	13	US-10-027-632-39522
3	22.4	67.9	639	13	US-10-027-632-39523
4	22.4	67.9	2815	10	US-09-312-762A-7
5	20.2	61.2	3059	12	US-10-007-926A-408
6	20	60.6	916	14	US-10-198-846-4325
7	19.8	60.0	3207	12	US-10-274-583-13
8	19.8	60.0	3227	12	US-10-274-583-9
9	19.8	60.0	3227	12	US-10-274-583-10
10	19.8	60.0	3679	10	US-09-909-320-244
11	19.8	60.0	3679	10	US-09-909-088B-244
12	19.8	60.0	3679	10	US-09-905-291A-244
13	19.8	60.0	3679	10	US-09-902-853-244
14	19.8	60.0	3679	10	US-09-907-824-244
15	19.8	60.0	3679	10	US-09-907-841-244
16	19.8	60.0	3679	11	US-09-904-011-244

C 17	19.8	60.0	3679	11	US-09-906-742-244	Sequence 244, App
C 18	19.8	60.0	3679	11	US-09-906-838-244	Sequence 244, App
C 19	19.8	60.0	3679	11	US-09-907-613-244	Sequence 244, App
C 20	19.8	60.0	3679	11	US-09-907-942-244	Sequence 244, App
C 21	19.8	60.0	3679	11	US-09-904-859-244	Sequence 244, App
C 22	19.8	60.0	3679	11	US-09-909-204-244	Sequence 244, App
C 23	19.8	60.0	3679	11	US-09-904-820-244	Sequence 244, App
C 24	19.8	60.0	3679	11	US-09-904-786-244	Sequence 244, App
C 25	19.8	60.0	3679	11	US-09-906-446-244	Sequence 244, App
C 26	19.8	60.0	3679	11	US-09-906-700-244	Sequence 244, App
C 27	19.8	60.0	3679	11	US-09-903-786-244	Sequence 244, App
C 28	19.8	60.0	3679	11	US-09-902-903-244	Sequence 244, App
C 29	19.8	60.0	3679	11	US-09-903-749A-244	Sequence 244, App
C 30	19.8	60.0	3679	11	US-09-904-119-244	Sequence 244, App
C 31	19.8	60.0	3679	11	US-09-904-956-244	Sequence 244, App
C 32	19.8	60.0	3679	11	US-09-902-736-244	Sequence 244, App
C 33	19.8	60.0	3679	11	US-09-907-794-244	Sequence 244, App
C 34	19.8	60.0	3679	11	US-09-903-943-244	Sequence 244, App
C 35	19.8	60.0	3679	11	US-09-904-562-244	Sequence 244, App
C 36	19.8	60.0	3679	11	US-09-907-925-244	Sequence 244, App
C 37	19.8	60.0	3679	11	US-09-902-692-244	Sequence 244, App
C 38	19.8	60.0	3679	11	US-09-903-520-244	Sequence 244, App
C 39	19.8	60.0	3679	11	US-09-905-056-244	Sequence 244, App
C 40	19.8	60.0	3679	11	US-09-909-064-244	Sequence 244, App
C 41	19.8	60.0	3679	11	US-09-904-553-244	Sequence 244, App
C 42	19.8	60.0	3679	11	US-09-905-381-244	Sequence 244, App
C 43	19.8	60.0	3679	11	US-09-905-088-244	Sequence 244, App
C 44	19.8	60.0	3679	11	US-09-907-575-244	Sequence 244, App
C 45	19.8	60.0	3679	11	US-09-905-075-244	Sequence 244, App

ALIGNMENTS

RESULT 1
US-09-827-688-8
; Sequence 8, Application US/09827688
; Publication No. US20030165476A1
; GENERAL INFORMATION:
; APPLICANT: ORSON, FRANK
; APPLICANT: KINSEY, BERRA
; APPLICANT: BROGAL, BALBIR
; TITLE OF INVENTION: MACROAGGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION 1
; TITLE OF INVENTION: AGENTS
; FILE REFERENCE: P01948U1/10004014
; CURRENT APPLICATION NUMBER: US/09/827,688
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/195,680
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 154746
; TYPE: DNA
; ORGANISM: HERPESVIRUS 2
US-09-827-688-8

Query Match 68.5%; Score 22.6; DB 12; Length 154746;
Best Local Similarity 95.7%; Pred. No. 4.3;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAAACTGCTGCTCCAGCAGT 28
DB 54077 CAACTGCTGCTCCAGCAGT 54099

RESULT 2
US-10-027-632-39522
; Sequence 39522, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome

```
FILE REFERENCE: 108627.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39522
LENGTH: 639
TYPE: DNA
ORGANISM: Human
US-10-027-632-39522
```

```
Query Match      67.9%; Score 22.4; DB 13; Length 639;
Best Local Similarity 88.5%; Pred. No. 4.7;
Matches 23; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      4 TCGAACTGCTGCTCTCCAGCATGC 29
      |||||
DB      581 TCGAACTGCTGCTCTCCAGCATGC 606
```

```
RESULT 3
US-10-027-632-39523
Sequence 39523, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108627.129
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39523
LENGTH: 639
TYPE: DNA
ORGANISM: Human
US-10-027-632-39523
```

```
Query Match      67.9%; Score 22.4; DB 13; Length 639;
Best Local Similarity 88.5%; Pred. No. 4.7;
Matches 23; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      4 TCGAACTGCTGCTCTCCAGCATGC 29
      |||||
DB      581 TCGAACTGCTGCTCTCCAGCATGC 606
```

```
RESULT 4
US-09-312-762A-7/c
Sequence 7, Application US/09312762A
Patent No. US20020115069A1
GENERAL INFORMATION:
APPLICANT: MIA HOROWITZ ET AL.
TITLE OF INVENTION: EH DOMAIN CONTAINING GENES AND PROTEINS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSES: Mark M. Friedman c/o Anthony Castorina
STREET: 2001 Jefferson Davis Highway, Suite 207
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
COMPUTER: Twinhead* Slimnote-890TX
OPERATING SYSTEM: MS DOS version 6.2,
OPERATING SYSTEM: Windows version 3.11
SOFTWARE: Word for Windows version 2.0 converted to
SOFTWARE: an ASCII file
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/312,762A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/026,898
FILING DATE: 20 FEB 1998
ATTORNEY/AGENT INFORMATION:
NAME: Friedman, Mark M.
REGISTRATION NUMBER: 33,883
REFERENCE/DOCKET NUMBER: 916/10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 972-3-562553
TELEFAX: 972-3-562554
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2815
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-312-762A-7
```

```
Query Match      67.9%; Score 22.4; DB 10; Length 2815;
Best Local Similarity 88.5%; Pred. No. 4.8;
Matches 23; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      4 TCGAACTGCTGCTCTCCAGCATGC 29
      |||||
DB      980 TCGAACTGCTGCTCTCCAGCATGC 955
```

```
RESULT 5
US-10-007-926A-408
Sequence 408, Application US/10007926A
Publication No. US20030143539A1
GENERAL INFORMATION:
APPLICANT: BERTUCCI, FRANCOIS
APPLICANT: HOUIGATTE, REMI
APPLICANT: BIRNBAUM, DANIEL
APPLICANT: NGUYEN, CATHERINE
APPLICANT: VIENS, PATRICE
APPLICANT: FERT, VINCENT
TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
FILE REFERENCE: 1546-R-00
CURRENT APPLICATION NUMBER: US/10/007,926A
CURRENT FILING DATE: 2001-12-07
PRIOR APPLICATION NUMBER: 60/254,090
```


PRIOR FILING DATE: 2000-12-08
NUMBER OF SEQ ID NOS: 468
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 408
LENGTH: 3059
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Interleukin enhancer binding factor 1 (ILF1)
US-10-007-926A-408

Query Match 61.2%; Score 20.2; DB 12; Length 3059;
Best Local Similarity 75.8%; Pred. No. 42;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CCTGCAAACTCGTCTCTCCAGTCAGG 33
Db 797 CGTGCAAACTCTGCTCCAGCCCGGAGG 829

RESULT 6
US-10-198-846-4325

Sequence 4325, Application US/10198846
Publication No. US2003009974A1
GENERAL INFORMATION:

APPLICANT: Lillie, James

APPLICANT: Xu, Yongyao

APPLICANT: Wang, Youzhen

APPLICANT: Steinmann, Kathleen

TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

FILE REFERENCE: Wf-049

CURRENT APPLICATION NUMBER: US/10/198,846

CURRENT FILING DATE: 2002-07-18

PRIOR APPLICATION NUMBER: 60/306,220

PRIOR FILING DATE: 2001-07-18

NUMBER OF SEQ ID NOS: 14084

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4325

LENGTH: 916

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

LOCATION: 439, 451, 525, 571, 596, 627, 636, 642, 668, 686, 689, 706,

LOCATION: 711, 715, 740, 759, 779, 782, 783, 789, 790, 796, 798, 802,

LOCATION: 812, 816, 817, 826, 828, 831, 835, 836, 837, 841, 842, 846,

LOCATION: 850, 856, 871, 872, 873, 876, 879, 880, 882, 883, 889

OTHER INFORMATION: n = A,T,C or G

FEATURE:

NAME/KEY: misc.feature

LOCATION: 890, 895, 896, 897, 898, 900, 901, 903, 904, 909, 910, 911,

LOCATION: 914

OTHER INFORMATION: n = A,T,C or G

US-10-198-846-4325

Query Match 60.6%; Score 20; DB 14; Length 916;
Best Local Similarity 76.7%; Pred. No. 49;
Matches 23; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 4 TGCAAACTCGTCTCTCCAGTCAGG 33
Db 486 TGCAAACTCGTCTCCAAACATCTTGG 515

RESULT 7
US-10-274-583-13/c

Sequence 13, Application US/10274583
Publication No. US20030138431A1
GENERAL INFORMATION:

TITLE OF INVENTION: LRCAPS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

FILE REFERENCE: EX02-119C

CURRENT APPLICATION NUMBER: US/10/274,583

CURRENT FILING DATE: 2002-10-21

PRIOR APPLICATION NUMBER: 60/338,733

PRIOR FILING DATE: 2001-10-22

PRIOR APPLICATION NUMBER: 60/357,600

PRIOR FILING DATE: 2002-02-15

PRIOR APPLICATION NUMBER: 60/361,196

PRIOR FILING DATE: 2002-03-01

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn version 3.1

SEQ ID NO 13

LENGTH: 3207

TYPE: DNA

ORGANISM: Homo sapiens

US-10-274-583-13

Query Match 60.0%; Score 19.8; DB 12; Length 3207;
Best Local Similarity 84.0%; Pred. No. 61;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCTGCAAACTCGTCTCTCCAGC 25
Db 793 CCTGCAAACTCGTCTCTCCAGC 769

RESULT 8
US-10-274-583-9/c

Sequence 9, Application US/10274583
Publication No. US20030138431A1
GENERAL INFORMATION:

APPLICANT: Exelixis, Inc.

TITLE OF INVENTION: LRCAPS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

FILE REFERENCE: EX02-119C

CURRENT APPLICATION NUMBER: US/10/274,583

CURRENT FILING DATE: 2002-10-21

PRIOR APPLICATION NUMBER: 60/338,733

PRIOR FILING DATE: 2001-10-22

PRIOR APPLICATION NUMBER: 60/357,600

PRIOR FILING DATE: 2002-02-15

PRIOR APPLICATION NUMBER: 60/361,196

PRIOR FILING DATE: 2002-03-01

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn version 3.1

SEQ ID NO 9

LENGTH: 3227

TYPE: DNA

ORGANISM: Homo sapiens

US-10-274-583-9

Query Match 60.0%; Score 19.8; DB 12; Length 3227;
Best Local Similarity 84.0%; Pred. No. 61;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCTGCAAACTCGTCTCTCCAGC 25
Db 805 CCTGCAAACTCGTCTCTCCAGC 781

RESULT 9
US-10-274-583-10/c

Sequence 10, Application US/10274583
Publication No. US20030138431A1
GENERAL INFORMATION:

APPLICANT: Exelixis, Inc.

TITLE OF INVENTION: LRCAPS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

FILE REFERENCE: EX02-119C

CURRENT APPLICATION NUMBER: US/10/274,583

CURRENT FILING DATE: 2002-10-21

PRIOR APPLICATION NUMBER: 60/338,733

PRIOR FILING DATE: 2002-02-15

PRIOR APPLICATION NUMBER: 60/361,196

PRIOR FILING DATE: 2002-03-01

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn version 3.1

SEQ ID NO 13

LENGTH: 3207

TYPE: DNA

ORGANISM: Homo sapiens

US-10-274-583-13

```
/ PRIOR FILING DATE: 2002-02-15
/ PRIOR APPLICATION NUMBER: 60/161,196
/ PRIOR FILING DATE: 2002-03-01
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 10
/ LENGTH: 3227
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-274-583-10

Query Match      60.0%; Score 19.8; DB 12; Length 3227;
Best Local Similarity 84.0%; Pred. No. 61;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1  CCTGCAAACTGCTGCTCTCCAGC 25
Db      805  CCTGCAAACTGCTGCTCTCCAGC 781

RESULT 10
US-09-909-320-244/C
/ Sequence 244, Application US/09909320
/ Patent No. US20020132240A1
/ GENERAL INFORMATION:
/ APPLICANT: Genentech, Inc.
/ APPLICANT: Ashkenazi, Avi
/ APPLICANT: Borstein, David
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Eaton, Dan L.
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerber, Hanspeter
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, A.
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, Christopher J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Hillan, Kenneth, J.
/ APPLICANT: Kijavlin, Ivar J.
/ APPLICANT: Mather, Jennie P.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William, I.
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
/ FILE REFERENCE: 10466-14
/ CURRENT APPLICATION NUMBER: US/09/909,320
/ CURRENT FILING DATE: 2002-01-04
/ PRIOR APPLICATION NUMBER: PCT/US00/04414
/ PRIOR FILING DATE: 2000-02-22
/ PRIOR APPLICATION NUMBER: US 60/143,048
/ PRIOR FILING DATE: 1999-07-07
/ PRIOR APPLICATION NUMBER: US 60/145,698
/ PRIOR FILING DATE: 1999-07-26
/ PRIOR APPLICATION NUMBER: US 60/146,222
/ PRIOR FILING DATE: 1999-07-28
/ PRIOR APPLICATION NUMBER: PCT/US99/20594
/ PRIOR FILING DATE: 1999-09-08
/ PRIOR APPLICATION NUMBER: PCT/US99/20944
/ PRIOR FILING DATE: 1999-09-13
/ PRIOR APPLICATION NUMBER: PCT/US99/21090
/ PRIOR FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: PCT/US99/21547
/ PRIOR FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: PCT/US99/23089
/ PRIOR FILING DATE: 1999-10-05
```

```
/ PRIOR APPLICATION NUMBER: PCT/US99/28214
/ PRIOR FILING DATE: 1999-11-29
/ PRIOR APPLICATION NUMBER: PCT/US99/28313
/ PRIOR FILING DATE: 1999-11-30
/ PRIOR APPLICATION NUMBER: PCT/US99/28564
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: PCT/US99/28565
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: PCT/US99/30095
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: PCT/US99/30911
/ PRIOR FILING DATE: 1999-12-20
/ PRIOR APPLICATION NUMBER: PCT/US99/30999
/ PRIOR FILING DATE: 1999-12-20
/ PRIOR APPLICATION NUMBER: PCT/US00/00219
/ PRIOR FILING DATE: 2000-01-05
/ NUMBER OF SEQ ID NOS: 423
/ SEQ ID NO 244
/ LENGTH: 3679
/ TYPE: DNA
/ ORGANISM: Homo Sapien
US-09-909-320-244

Query Match      60.0%; Score 19.8; DB 10; Length 3679;
Best Local Similarity 84.0%; Pred. No. 62;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1  CCTGCAAACTGCTGCTCTCCAGC 25
Db      1297  CCTGCAAACTGCTGCTCTCCAGC 1273

RESULT 11
US-09-909-088B-244/C
/ Sequence 244, Application US/09909088B
/ Patent No. US20020146709A1
/ GENERAL INFORMATION:
/ APPLICANT: Genentech, Inc.
/ APPLICANT: Ashkenazi, Avi
/ APPLICANT: Borstein, David
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Eaton, Dan L.
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerber, Hanspeter
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, A.
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, Christopher J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Hillan, Kenneth, J.
/ APPLICANT: Kijavlin, Ivar J.
/ APPLICANT: Mather, Jennie P.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William, I.
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
/ FILE REFERENCE: 10466-14
/ CURRENT APPLICATION NUMBER: US/09/909,088B
/ CURRENT FILING DATE: 2001-07-18
/ PRIOR APPLICATION NUMBER: PCT/US00/04414
/ PRIOR FILING DATE: 2000-02-22
/ PRIOR APPLICATION NUMBER: US 60/143,048
/ PRIOR FILING DATE: 1999-07-07
/ PRIOR APPLICATION NUMBER: US 60/145,698
/ PRIOR FILING DATE: 1999-07-26
```

PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 244
LENGTH: 3679
TYPE: DNA
ORGANISM: Homo Sapien
US-09-909-0888-244

Query Match 60.0%; Score 19.8; DB 10; Length 3679;
Best Local Similarity 84.0%; Pred. No. 62;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Oy 1 CCTGCAGAACTCGTGTCTCTCAGC 25
Db 1297 CCTGCAGAACTCGTGTCTCTCAGC 1273

RESULT 12
US-09-905-291A-244/C
Sequence 244, Application US/09905291A
Patent No. US20020160374A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Grimaldi, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kiljavin, Ivar J.
APPLICANT: Macher, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905,291A
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 244
LENGTH: 3679
TYPE: DNA
ORGANISM: Homo Sapien
US-09-905-291A-244

Query Match 60.0%; Score 19.8; DB 10; Length 3679;
Best Local Similarity 84.0%; Pred. No. 62;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Oy 1 CCTGCAGAACTCGTGTCTCTCAGC 25
Db 1297 CCTGCAGAACTCGTGTCTCTCAGC 1273

RESULT 13
US-09-902-853-244/C
Sequence 244, Application US/09902853
Publication No. US20020192659A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, A.

APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Guiney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/902,853
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US/09/665,350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 244
LENGTH: 3679
TYPE: DNA
ORGANISM: Homo Sapien
US-09-902-853-244

Query Match 60.0%; Score 19.8; DB 10; Length 3679;
Best Local Similarity 84.0%; Pred. No. 62;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCTGCAACTGCTGCTCTCCAGC 25
DB 1297 CCTGCAAGCTGCTGCTCTCCAGC 1273

RESULT 14
US-09-907-824-244/C
Sequence 244, Application US/09907824
Publication No. US20020197671A1
GENERAL INFORMATION:

APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Bockstein, David
APPLICANT: Deenoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Geo, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Guiney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/907,824
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: 09/665,350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 244
LENGTH: 3679
TYPE: DNA
ORGANISM: Homo Sapien
US-09-907-824-244

Query Match 60.0%; Score 19.8; DB 10; Length 3679;
Best Local Similarity 84.0%; Pred. No. 62;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCTGCAAACTCGTGTCTCCAGC 25
Db 1297 CCTGCAAACTCGTGTCTCCAGC 1273

RESULT 15

US-09-907-841-244/C
Sequence 244, Application US/09907841
Publication No. US20020198366A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.
APPLICANT: Askenazi, Avi
APPLICANT: Borstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlisen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijaviri, Ivar J.
APPLICANT: Macher, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 1046-14
CURRENT APPLICATION NUMBER: US/09/907,841
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 244
LENGTH: 3679

TYPE: DNA

ORGANISM: Homo Sapien

US-09-907-841-244

Query Match 60.0%; Score 19.8; DB 10; Length 3679;

Best Local Similarity 84.0%; Pred. No. 62;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCTGCAAACTCGTGTCTCCAGC 25
Db 1297 CCTGCAAACTCGTGTCTCCAGC 1273

Search completed: September 16, 2003, 22:54:39
Job time : 120.257 secs

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 18:22:25 ; Search time 1670.26 Seconds
(without alignment)
480.194 Million cell updates/sec

Title: US-09-594-065-4

Perfect score: 33
Sequence: 1 cccgcgaactcgtctcctccagatgcaggg 33

Scoring table: IDENTITY NUC
Gapop 10-0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
C 1	23.6	71.5	832 28	BH049615 RPCI-24-3
C 2	23.6	71.5	3863 11	AK030462 Mus muscu
C 3	22.4	67.9	590 14	CB159107 K-EST0218
C 4	22.4	67.9	605 9	AA972967 op25d11.s

Result No.	Score	Query Match length	ID	Description
C 5	22.4	67.9	639 2	BSM096537
C 6	22.4	67.9	639 14	CB155818
C 7	22.4	67.9	884 14	CA488926
C 8	22.2	67.3	759 29	BZ375631
C 9	22.2	66.7	401 13	BY540314
C 10	22.2	66.7	405 13	BY541787
C 11	22.2	66.7	439 13	BY536425
C 12	22.2	66.7	463 13	BY535446
C 13	22.2	66.7	651 12	B1528489
C 14	22.2	66.7	651 12	B1528424
C 15	21.8	66.1	398 28	AO587305
C 16	21.4	64.8	854 12	B1602858
C 17	21.2	64.2	476 28	AQ527090
C 18	21.2	64.2	866 14	CD0303466
C 19	20.8	63.0	299 9	AA069603
C 20	20.8	63.0	445 10	BE095803
C 21	20.8	63.0	500 9	AM533901
C 22	20.8	63.0	554 12	BM721965
C 23	20.8	63.0	566 13	BQ779947
C 24	20.8	63.0	593 10	BF395171
C 25	20.8	63.0	614 10	AM915165
C 26	20.8	63.0	649 12	BM702410
C 27	20.8	63.0	670 14	CA339977
C 28	20.8	63.0	996 13	BQ716338
C 29	20.8	63.0	1010 29	CC026919
C 30	20.6	62.4	2203 11	BC022928
C 31	20.6	62.4	552 10	BM237104
C 32	20.4	61.8	491 10	BF622435
C 33	20.4	61.8	522 9	AL632202
C 34	20.4	61.8	593 14	CA360715
C 35	20.4	61.8	608 28	BH586809
C 36	20.4	61.8	624 28	BH741731
C 37	20.4	61.8	628 13	BX309994
C 38	20.4	61.8	681 28	BH886140
C 39	20.4	61.8	691 29	AG080307
C 40	20.4	61.8	725 13	BU260794
C 41	20.4	61.8	759 29	BZ599488
C 42	20.4	61.8	780 13	BU459650
C 43	20.4	61.8	913 13	BU169492
C 44	20.2	61.2	157 13	BY316884
C 45	20.2	61.2	354 13	BY034960

ALIGNMENTS

RESULT 1
BH049615/c
LOCUS
DEFINITION
RPCI-24-319P13_TV RPCI-24 Mus musculus genomic clone RPCI-24-319P13
ACCESSION
BH049615
VERSION
BH049615.1 GI:14838781
KEYWORDS
SOURCE
MUS musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 832)
Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akincet,B., Levins,M.,
Teegaye,G., Geer,K., Krol,M., Shartsbeyn,A., Gebregorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-24
JOURNAL
Unpublished
Other_GSSs: RPCI-24-319P13_TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhaoc@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC

library availability, please contact Pieter de Jong
(pdjong@mail.chc.org). Clones may be purchased from BACPAC
Resources (<http://www.chori.org/bacpac/orderingframe.html>). BAC end
pages: http://www.tigr.org/tdb/bac_end/mouse/bac_end_intro.html
Placer: 319 row: P column: 13
Seq primer: SPE
Class: BAC ends.

FEATURES

source
Location/Qualifiers
1..832
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPPI-24-319F13"
/sex="Male"
/cell_type="Spleen/Brint"
/clone_lib="RPPI-24"
/note="Vector: pTARBAC1, Site 1: BamHI; Site 2: BamHI;
RPPI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

BASE COUNT 238 a 147 c 178 g 269 t

ORIGIN

Query Match 71.5%; Score 23.6; DB 28; Length 832;
Bent Local Similarity 81.2%; Pred. No. 1.8e+02;
Matches 26; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CCCTGGAACTGCTGCTCTCCACGATCGAGG 32
110 CCCAGCAACTGCTGCTACCAACATCGAGG 79

RESULT 2

LOCUS AK030462 3863 bp mRNA linear HTC 05-DEC-2002
DEFINITION Mus musculus adult male pituitary gland cDNA, RIKEN full-length
enriched library, clone:5330415G09 product:unknown EST, full insert
sequence.

ACCESSION AK030462.1 GI:26081833
VERSION AK030462.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

JOURNAL High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

REFERENCE 3
10349636

REFERENCE 4
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE 5
20499374
11042159

REFERENCE 6
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Chata, S., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

REFERENCE 7
20530913

JOURNAL misc_feature

REFERENCES

11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, T., Hara, A., Kikunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I.,
Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaio, T., Pesole, G.,
Quackenbush, J., Schriml, L.M., Stambli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsi, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Meshima, J., Mazzarelli, J., Mombere, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohno, T., S.
and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

JOURNAL MEDLINE 21085660
PUBMED 11217851

TITLE

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

6 (bases 1 to 3863)

Adachi, J., Aizawa, K., Hanagaki, T., Hara, A., Hashizume, W.,
Fukuda, S., Furuno, M., Hayashizaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozawa, T.,
Hori, P., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Konda, M.,
Koye, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Nishizaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, R., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (16-JUN-2001) Yoshitake Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Science Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel:81-45-503-9222,
Fax:81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Science Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.

LOCATION/Qualifiers
1..3863

FEATURES
source
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM,DB:5330415G09"
/db_xref="taxon:10090"
/clone="5330415G09"
/sex="male"
/cell_type="pituitary gland"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

1..3863

misc_feature


```

/ncseq="unknown EST (GB|BE334066, evidence: BLASTN, 97%
match=403)"
BASE COUNT      1159 a      794 c      778 g      1132 t
ORIGIN

```

Query Match	71.5%	Score 23.6	DB 11	length 3863;
Best Local Similarity	81.2%	Pred. No. 2.8e+02;		
Matches	26;	Conservative	1;	Mismatches 5;
				Indels 0;
				Gaps 0

0y 1 CCCGCAAACTCGTGTCTCCAGATGCGG 32
 1433 CCCAGCAAACTCGTGTACACAAACATGCGG 1464
 Db

LOCUS	590 bp	mRNA	linear	EST
CB159107	590 bp	mRNA	linear	EST 30-JAN-2003
CB159107/c	590 bp	mRNA	linear	EST 30-JAN-2003
RESULT 3				

5', mRNA sequence.

KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (Bases to 590)
Kim,N.S., Hahn,Y., Oh,J.H.,Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

TITLE	21C Frontier Korean EST Project 2001
JOURNAL	Unpublished
COMMENT	Contact: Kim YS

Dr. Heungsik Lee, Associate Professor of Biochemistry
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409

```

FEATURES
source      1 590
File: 37 LOW: 5 column: 05
High quality sequence stop: 590.
Location/Qualifiers

```

```

/col_type="mRNA"
/db_xref="taxon:3606"
/clone="L18POL1n1-37-E05"
cell:1;flow:354;chip:

```

```
/clone_lib="libPOOLint"  
/note="Organ: Liver; Vector: pT73-Pac; Site_1: EcoRI;  
Site_2: NotI; The library was contributed by the Soares,
```

BASE COUNT	6 (9) : 791-806. RNA was prepared from harvested cell culture."	146 a	158 c	187 g	99 t
------------	---	-------	-------	-------	------

Query Match	67.9%	Score 22.4	DB 14	Length 590
Best Local Similarity	88.5%	Pred. No. 4.5e+02		
Matches 23; Conservative	1	Mismatches 2	Indels 0	Gaps 0

Db 359 TGCAACTGTCGTCTCCAGCATGC 334

[illegible]

ACCESSION	AA972967	
VERSION	AA972967.1	GI:3148147
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
1 (bases 1 to 605)
AUTHORS
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

COMMENT.....
 ..cgsapsb@rovaly-fre.com
 Contact: Robert Strausberg, Ph.D.
 Email: cgsapsb-r@mail.nih.gov
 This clone is available rovaly-free through LWN! : contact the

FEATURES

- AmerSham Binding: 932 Glc Biol: 0.00
- Seq primer: ~40ml3 fwd. ET from AmerSham
- High quality sequence scop: 444.
- location/Omalifera

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NMGR.157807"

```

/Clone_lib="Soares_NFL_T_GBC_S1"
 /note="Organ: pooled; Vector: pTR73D-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not I; Site 2: Eco RI;
 Equal amounts of plasmid DNA from three recombinized

NCI OGAP GCBI) were mixed, and ss circles were made *in vitro*. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver

I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Patricia Bonaldo."

Query Match	67.9%	Score 22.4;	DB 9;	Length 605;
Best Local Similarity	88.5%;	Pred. No. 4.5e+02;		

Oy 4 TGCAACTCGTGTCTCCTCCAGCATGC 29
Ox :
222 TGCAACTCGTGTCTCCTCCAGCATGC 247

RESULT 5
HSM096537/c
ID HSM096537 standard; RNA: 639 BP.

MC BA302820.1
XX
XX
SV BX502820.1
XX

XX Homo sapiens mRNA; EST DKFZp779B0170_r1 (from clone DKFZp779B0170)
DE
XX
DI 05-Mar-2003 (rel. 7.5, last updated, version 1)
DI

AA Homo sapiens (human)
OS Eukaryota; Metazoa;
OC Chordata; Vertebrata; Euteleostomi; Mammalia
Phyla: Deuterostoma; Cephalochordata; Vertebrata; Mammalia

RM 11
RP 1-639

cancer genes encoding membrane and secreted proteins.

Manuscript submitted."

BASE COUNT 238 a 231 c 216 g 199 t

ORIGIN

Query Match 67.9% Score 22.4; DB 14; Length 884;
Best Local Similarity 88.5%; Pred. No. 5.1e+02;
Matches 23; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

4 TGCAACTGCTGCTCCTCCAGCATGC 29
|||||||:|||||||
54 TGCAACTGCTGCTCCTCCAGCATGC 29

RESULT 8
LOCUS B2375631

DEFINITION B2375631 759 bp DNA linear GSS 26-NOV-2002
1659907.5 genomic survey sequence.

ACCESSION B2375631
VERSION B2375631.1 GI:25463412
KEYWORDS GSS.

SOURCE

ORGANISM

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dechhia, N.,
Katzendburger, P., King, L., Miller, B., Muller, S., Nascimben, L.,
Zutavern, T., McCombie, W.R. and Martienssen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)

TITLE

JOURNAL

Unpublished

Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: 1659 row: 5 column: 07
Seq primer: -21M13UniRev
Class: shotgun
High quality sequence stop: 759.

Location/Qualifiers

FEATURES

SOURCE

1..759
/organism="Zea mays"
/mol_type="genomic DNA"
/cultiivar="B73"
/db_xref="taxon:4577"
/clone="1659907"
/lab_host="DH5a"
/clone_lib="WGS-ZmaysF (DH5a methyl filtered)"
/note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector (x/y
reads in M13mp19, b/g reads in pUC19). The same ligation
was transformed into DH5a."

BASE COUNT 199 a 174 c 175 g 211 t

ORIGIN

Query Match 67.3% Score 22.2; DB 29; Length 759;
Best Local Similarity 82.8%; Pred. No. 5.7e+02;
Matches 24; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

2 CCTGCAACTGCTGCTCCTCCAGCATGCA 30
|||||||:|||||||
542 CCTGCAACTGCTGCTCCTCCAGCATGCA 570

RESULT 9

BY540314/C

LOCUS BY540314

DEFINITION BY540314 RIKEN full-length enriched, B6-derived C01 +ve dendritic
cells Mus musculus cDNA clone F730006E21.3, mRNA sequence.
BY540314
BY540314.1 GI:26874693
EST.

ACCESSION BY540314.1 GI:26874693

VERSION BY540314

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 401)
Okazaki, Y., Furuno, M., Kasekawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oseko, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D.P., Bul, C., Hume, D.A.,
Quackenbush, J., Schram, L.M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V., Chochia, C., Corbani,
L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,
A., Frazer, K.S., Gasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimond, S., Guerinich, S., Hirokawa, N., Jackson, I.J.,
Jarvis, E.D., Kanai, A., Kawai, H., Kawasawa, Y., Kezierski, R.M.,
King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Leinhardt, B., Lyons,
P.A., Maglocz, D.R., Maltas, K., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,
Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring,
B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou,
M., Shimada, K., Sultana, R., Takekura, Y., Taylor, M.S., Teasdale,
R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L.G., Wyszewski, A., Yamaoka, A.,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura,
M., Sakazume, N., Seto, K., Shiraki, T., Maki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, I., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E.S., Rogers, J., Birney, R. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT

Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Aikawa, T.,
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A.,
Mura, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
M., Maki, K., Watanabe, A., Watanabe, M. and Hayashizaki, Y. Direct
Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/RNC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source location/Qualifiers
1..401
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F73006E21"
/cell_type="B6-derived CD11 +ve dendritic cells"
/clone_11b="RIKEN full-length enriched, B6-derived CD11 +ve dendritic cells"
BASE COUNT 90 a 105 c 106 g 100 t
ORIGIN

Query Match 66.7%; Score 22; DB 13; Length 401;
Best Local Similarity 78.1%; Pred. No. 5,6e+02;
Matches 25; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CCTGCAAACTGTCCTCCGACATGCAG 32
37 CCTGCAATCTCTGTCCTCTCTCGGACAG 6

RESULT 10 405 bp mRNA linear EST 14-DEC-2002
LOCUS BY541787/c
DEFINITION BY541787 RIKEN full-length enriched, B6-derived CD11 +ve dendritic cells Mus musculus cDNA clone F730022F15 3', mRNA sequence.
ACCESSION BY541787
VERSION BY541787.1 GI:26876166
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 405)
Otakeki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bone, H., Kondo, S., Nikaido, I., Ogata, N., Saito, R., Suzuki, H., Yamazaki, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schorbach, C., Gojodori, I., Balderelli, R., Hill, D. P., Bull, C., Hume, D. A., Queckenbush, J., Schrieml, L. M., Kanapin, A., Matluda, H., Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Brusc, V., Chotila, C., Corbett, L. B., Couvine, S., Datta, E., Dragan, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Glas, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hinkawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedietsaki, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Piliat, R., Pontius, J. U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempel, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wallestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilm, L. G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carlinici, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Aikawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sakai, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Watanabe, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL MEDLINE 22354683

PUBMED COMMENT

12466851
Contact: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Shukho-cho, Tsukuba, Ibaraki, 305-8575, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp
URL: <http://genome.gsc.riken.go.jp/>

Alizawa, K., Akimura, T., Aizawa, T., Carlinici, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Watanabe, M., and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichannel sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/RNC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source location/Qualifiers
1..405
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F730022F15"
/cell_type="B6-derived CD11 +ve dendritic cells"
/clone_11b="RIKEN full-length enriched, B6-derived CD11 +ve dendritic cells"
BASE COUNT 92 a 107 c 106 g 100 t
ORIGIN

Query Match 66.7%; Score 22; DB 13; Length 405;
Best Local Similarity 78.1%; Pred. No. 5,7e+02;
Matches 25; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CCTGCAAACTGTCCTCCGACATGCAG 32
37 CCTGCAATCTCTGTCCTCTCTCGGACAG 6

RESULT 11 439 bp mRNA linear EST 14-DEC-2002
LOCUS BY536425/c
DEFINITION BY536425 RIKEN full-length enriched, NOD-derived CD11 +ve dendritic cells Mus musculus cDNA clone F630310H17 3', mRNA sequence.
ACCESSION BY536425
VERSION BY536425.1 GI:26870804
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS

REFERENCE 1 (bases 1 to 439)
 AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishida, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bull, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Belser, K.W., Blake, J.A., Brad, D., Brusc, V., Chotila, C., Corbett, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Guernich, S., Hirokawa, N., Jackson, J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzielski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perle, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Saito, K., Schneider, C., Sempke, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawaji, J., Alizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasuniishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 12466851
 22354683
 JOURNAL MEDLINE PUBMED
 COMMENT
 TITLE
 CONTACT: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-resgsc.riken.go.jp/
 URL: http://genome-gsc.riken.go.jp/
 Aikawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichannel sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC Building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
 Please visit our web site (http://genome-gsc.riken.go.jp) for further details.
 FEATURES
 SOURCE
 1. 439
 /organism="Mus musculus"
 /mol_type="mRNA"

BASE COUNT 97 a 116 c 122 g 104 t
 ORIGIN
 Query Match 66.7% Score 22; DB 13; Length 439;
 Best Local Similarity 78.1% Pred. NO. 5.8e+02;
 Matches 25; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CCTGCAAACTGCTGCTCTCCAGCATGACG 32
 Db 78 CCTGCAAACTGCTGCTCTCCAGCATGACG 47
 RESULT 12
 LOCUS BY535446/C
 DEFINITION
 BY535446 RIKEN full-length enriched, NOD-derived CD1c +ve dendritic cells Mus musculus cDNA clone F630305G21 3', mRNA
 ACCESSION
 VERSION BY535446
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 REFERENCE
 AUTHORS
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishida, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bull, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Belser, K.W., Blake, J.A., Brad, D., Brusc, V., Chotila, C., Corbett, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Guernich, S., Hirokawa, N., Jackson, J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzielski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perle, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Saito, K., Schneider, C., Sempke, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawaji, J., Alizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasuniishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 12466851
 22354683
 JOURNAL MEDLINE PUBMED
 COMMENT
 TITLE
 CONTACT: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-resgsc.riken.go.jp/
 URL: http://genome-gsc.riken.go.jp/
 Aikawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane

T., Inotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct Substantiation

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome, 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedias: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC Building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

1.463
location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="taxon:10090"
/clone="P630305G21"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched, NOD-derived CD11c +ve dendritic cells"
BASE COUNT 102 a 120 c 131 g 110 t

Query Match 66.7%; Score 22; DB 13; Length 463;
Best Local Similarity 78.1%; Pred. No. 5.9e+02;
Matches 25; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCTGCAACTCGTGTCTCCAGCATGCGAGG 32
Db 96 CCTGCACTCTGTGTCTCTGCGGCGAGAGG 65

RESULT 13
B1528489 643 bp mRNA linear EST 29-AUG-2001
LOCUS 1024091A09.y1 C. reinhardtii CC-1690, normalized, Lambda Zap II
DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION B1528489
VERSION B1528489.1 GI:15369063
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 643)
Grosman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1024b
Unpublished
JOURNAL Contact: Charles Hauser
COMMENT DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159

Fax: 919 613 8177
Email: chauser@duke.edu
location/Qualifiers
1.463
/organism="Chlamydomonas reinhardtii"

FEATURES

source

/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap II"

/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO₂ and HS medium bubbled with 5% CO₂. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into Lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the Lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 126 a 178 c 226 g 113 t

Query Match 66.7%; Score 22; DB 12; Length 643;
Best Local Similarity 78.1%; Pred. No. 6.5e+02;
Matches 25; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 CCTGCAACTCGTGTCTCCAGCATGCGAGG 33
Db 112 CATTGCAACCGCTGTCTCTCAGCATGCGAGG 143

RESULT 14
B1528424 651 bp mRNA linear EST 29-AUG-2001
LOCUS 1024090E05.y1 C. reinhardtii CC-1690, normalized, Lambda Zap II
DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION B1528424
VERSION B1528424.1 GI:15368998
KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 651)
Grosman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.

Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1024b
Unpublished
JOURNAL Contact: Charles Hauser
COMMENT DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu

FEATURES

source

1.461
location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the

light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO₂ and HS medium bubbled with 5% CO₂. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with Exsistat (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 127 a 179 c 230 g 113 t 2 others

Query Match 66.7%; Score 22; DB 12; Length 651;
Best Local Similarity 78.1%; Pred. No. 6.5e+02;
Matches 25; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

CY 2 CCTGCAAACTGCTGCTCTCCAGCATGACAGG 33
DB 112 CGTGCACACCCGTTGCTCTCCAGCATGACAGG 143

RESULT 15
LOCUS AO587305/c 398 bp DNA linear GSS 07-JUN-1999
DEFINITION RPCI-11-452E3.TU RPCI-11 Homo sapiens genomic clone RPCI-11-452E3,
genomic survey sequence.

ACCESSION AO587305

VERSION AO587305.1 GI:5013985

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Zhao,S., Adams,M.D., Nieman,W., Malek,J., de Jong,P. and Venter
J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building

JOURNAL Unpublished

COMMENT Other GSSes: RPCI-11-452E3.TV

Contact: Shaying Zhao, William Nieman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@jlong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: SP6
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..398
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7673282"
/db_xref="taxon:9606"
/clone="RPCI-11-452E3"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"

BASE COUNT 73 a 117 c 98 g 101 t 9 others

Query Match 66.1%; Score 21.8; DB 28; Length 398;
Best Local Similarity 82.1%; Pred. No. 6.7e+02;
Matches 23; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

CY 1 CCTGCAAACTGCTGCTCTCCAGCATG 28
DB 104 CCTGCAACCCCGTTGCTCTCCAGCATG 77

Search completed: September 16, 2003, 20:40:01
Job time : 1674.26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 16, 2003, 17:35:29 ; Search time 657.921 Seconds

(without alignments)
1865.405 Million cell updates/sec

Title: US-09-594-065-3

Perfect score: 30

Sequence: 1 cccctgctgtagtgcgacgacctccgcaggg 30

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenBank: 1: gb_ba: 2: gb_bg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pac: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sfr: 12: gb_by: 13: gb_un: 14: gb_vl: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_in: 19: em_mu: 20: em_om: 21: em_or: 22: em_ov: 23: em_pac: 24: em_ph: 25: em_pl: 26: em_ro: 27: em_sfr: 28: em_un: 29: em_vl: 30: em_hg_hum: 31: em_hg_inv: 32: em_hg_other: 33: em_hg_mus: 34: em_hg_pln: 35: em_hg_rtd: 36: em_hg_mam: 37: em_hg_vrt: 38: em_ey: 39: em_hgo_hum: 40: em_hgo_mus: 41: em_hgo_other:

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20.6	68.7	40281	1	MLCB22
2	20.6	68.7	348450	1	MLCB22/c
3	20.4	68.0	11282	1	AB003979
4	20.4	68.0	302835	1	AB012555
5	20.2	67.3	110000	2	LMFCH34_04
6	19.4	64.7	68974	2	AC074063
7	19.4	64.7	110079	10	AC002315
8	19.4	64.7	163132	2	AC121312
9	19.4	64.7	176222	2	AC117138
10	19.4	64.7	194985	10	AC002406
11	19.4	64.7	200548	10	AL672026
12	19.4	64.7	217622	2	AC126569
13	19.4	64.7	222869	2	AC113777
14	19.4	64.7	226382	2	BX294168
15	19.4	64.7	226790	2	AC107570
16	19.4	64.7	229040	2	BX510361
17	19.4	64.7	231268	2	AC118318
18	19.4	64.7	241998	2	AC112739
19	19.4	64.7	263091	2	AC120822
20	19.4	64.7	267971	2	AC128955
21	19.4	64.7	270706	2	AC129862
22	19.4	64.7	299800	1	AP005040
23	19.2	64.0	182462	9	AC092669
24	19.2	63.3	94255	9	AC003090
25	19.2	63.3	110000	2	AC115691_2
26	19.2	63.3	122279	2	AC128646
27	19.2	63.3	146432	2	AC079358
28	19.2	63.3	166860	2	AL451002
29	19.2	63.3	182459	2	AC116818
30	19.2	63.3	197448	2	AC138342
31	19.2	63.3	200321	2	AC129934
32	19.2	63.3	207265	10	AC125157
33	19.2	63.3	211397	10	AC098887
34	18.8	62.7	576	6	AX122231
35	18.8	62.7	576	6	BD164348
36	18.8	62.7	702	9	HS434046
37	18.8	62.7	732	9	HS433685
38	18.8	62.7	751	9	HS434004
39	18.8	62.7	753	9	HS434007
40	18.8	62.7	793	9	HS434003
41	18.8	62.7	92019	9	AL450320
42	18.8	62.7	94244	9	AC063938
43	18.8	62.7	110000	2	AC001656_6
44	18.8	62.7	113454	2	AC017944
45	18.8	62.7	120871	9	AP001054

ALIGNMENTS

RESULT 1
LOCUS MLCB22/c
DEFINITION Mycobacterium leprae cosmid B22.
ACCESSION Z98741
VERSION Z98741.1 GI:2342602
KEYWORDS 30S ribosomal protein S15; ABC-type transporter; aminomethyl transferase; aminopeptidase; branched-chain amino acid aminotransferase; cobT; cybB; cykA; cysteine synthase; dihydrolipoamide succinyltransferase; gcvT; gspA; guanosine pentaphosphate synthetase; llyE; llyA; llyB; lipote-protein ligase; lipote acid synthetase; major membrane protein I; mmpI; multifunctional enzyme; nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase; nifs-like protein; oxidoreductase; pepA; polyribonucleotide

nucleotidyltransferase; protease/peptidase; pseudogene; rbf;
 riboflavin kinase; RLEP; RLEP2; rpoO; serine acetyltransferase;
 sucB.
 SOURCE
 ORGANISM
 Mycobacterium leprae
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
 REFERENCE
 AUTHORS
 1 (bases 1 to 40281)
 Eiglmeyer, K., Honore, N., Woods, S.A., Caudron, B. and Cole, S.T.
 Use of an ordered cosmid library to deduce the genomic organization
 of Mycobacterium leprae
 JOURNAL
 MEDLINE
 PUBMED
 8446027
 93188700
 Mol. Microbiol. 7 (2), 197-206 (1993)
 REFERENCE
 AUTHORS
 2 (bases 1 to 40281)
 Devlin, K. and Churcher, C.M.
 JOURNAL
 TITLE
 3 (bases 1 to 40281)
 Parkhill, J., Barrell, B.G. and Rajandream, M.A.
 Direct Submission
 Submitted (22-AUG-1997) Mycobacterium leprae sequencing project,
 Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
 CB10 1SA E-mail: barrell@sanger.ac.uk Cosmid supplied by Dr.
 Stewart T. Cole, (3) Unite de Genetique Moleculaire Bacterienne,
 Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15,
 France Requests for cosmids should be sent to Karin Eiglmeyer
 (ke@pasteur.fr)
 COMMENT
 Notes:
 The Sanger Centre is funded to complete the sequence of M. leprae
 by the Helser Program for Research in Leprosy and Tuberculosis of
 The New York Community Trust.
 Work in Paris is supported by the Helser Trust, the Association
 Francaise Raoul Follereau and the Groupement de Recherches et des
 Etudes des Genomes (GEP-GREG).
 Details of M. leprae sequencing at the Sanger Centre are available
 on the World Wide Web.
 (URL: <http://www.sanger.ac.uk/projects/>)
 CDS are numbered using the following system eg MLCB33.01c. ML (M.
 leprae), c333 (cosmid name), .01 (first CDS), c (complementary
 strand).
 The more significant matches with motifs in the PROSITE database
 are also included but some of these may be fortuitous. The length
 in codons is given for each CDS.
 Usually the highest scoring match found by fasta -o is given for
 CDS which show significant similarity to other CDS in the database.
 The position of possible ribosome binding site sequences are given
 where these have been used to deduce the initiation codon. All CDS
 over 100 codons have been analysed. Gene prediction is based on
 positional base preference in codons especially where there is an
 increase in the observed/expected third position G + C. CAUTION:
 We may not have predicted the correct initiation codon. Where
 possible we choose an initiation codon (atg, gtg, or ttg) which is
 preceded by an upstream ribosome binding site sequence (optimally
 5-13bp before the initiation codon). If this cannot be identified
 we choose the most upstream initiation codon.
 IMPORTANT: This sequence MAY NOT be the entire insert of the
 sequenced clone. It may be shorter because we only sequence
 overlapping sections once, or longer, because we arrange for a
 small overlap between neighbouring submissions. Cosmid B22 is
 overlapped by B151 at the 5' end, and by B129 at the 3' end.
 There are at least 9 conflicts between this sequence and the
 previously published sequences; in each case our sequence has been
 checked and is thought to be correct.
 FEATURES
 source
 1..40281
 location/Qualifiers
 /organism="Mycobacterium leprae"
 /mol_type="genomic DNA"
 /db_xref="taxon:1769"
 /clone="cosmid B22"
 complement(1..7427)
 /note="overlap with EM_BA_MLB151CS L78813 Mycobacterium
 leprae cosmid B1551 from: 2 to: 7425"
 complement(1..198)
 misc_feature
 gene
 CDS

CDS
 /gene="MLCB22.01c"
 complement(1..198)
 /gene="MLCB22.01c"
 /note="MLCB22.01c, unknown, partial CDS, len >66 aa;
 similar to hypothetical proteins from a range of organisms
 eg. YADR_ECOLI_P37026 hypothetical 12.1 kd protein in
 hemi-pfco intergenic region (114 aa), fasta scores: Opt:
 164 z-score: 323.9 E(): 6.3e-11, 45.1% identity in 51 aa
 overlap"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein MLCB22.01c"
 /protein_id="CAB11366.1"
 /db_xref="GI:2342603"
 /db_xref="SPTREMBL:O32951"
 /translation="MAVQNETLSAKTHGVILTDVATKASLLDQSGRDDLALRIAVQP
 GGCAGLRYNLFDDRTLDGDLT"
 complement(373..567)
 /note="MLCB22.02c, probable pseudogene fragment, len 64
 aa; similar to part of Y00G MYCTU Q10394 hypothetical 48.6
 kd protein cy190.16 (480 aa) (57.8% identity in 64 aa
 overlap)"
 /pseudo
 /codon_start=1
 /transl_table=11
 /db_xref="PID:e335004"
 /db_xref="PSEUDO:CAB11367.1"
 complement(566..670)
 /note="MLCB22.03c, probable pseudogene fragment, len: 34
 aa; similar to part of Y00G MYCTU Q10394 hypothetical 48.6
 kd protein cy190.16 (480 aa) (76.5% identity in 34 aa
 overlap)"
 /pseudo
 /codon_start=1
 /transl_table=11
 /db_xref="PID:e335005"
 /db_xref="PSEUDO:CAB11368.1"
 complement(799..861)
 /note="MLCB22.04c, probable pseudogene fragment, len: 20
 aa; similar to part of Y00G MYCTU Q10394 hypothetical 48.6
 kd protein cy190.16 (480 aa) (75.0% identity in 20 aa
 overlap)"
 /pseudo
 /codon_start=1
 /transl_table=11
 /db_xref="PID:e335006"
 /db_xref="PSEUDO:CAB11369.1"
 complement(1092..1181)
 /note="MLCB22.05c, probable pseudogene fragment, len: 29
 aa; similar to part of Y00G MYCTU Q10394 hypothetical 48.6
 kd protein cy190.16 (480 aa) (63.0% identity in 27 aa
 overlap)"
 /pseudo
 /codon_start=1
 /transl_table=11
 /db_xref="PID:e335007"
 /db_xref="PSEUDO:CAB11370.1"
 complement(1220..1312)
 /note="MLCB22.06c, probable pseudogene fragment, len: 30
 aa; similar to part of Y00G MYCTU Q10394 hypothetical 48.6
 kd protein cy190.16 (480 aa) (63.3% identity in 30 aa
 overlap)"
 /pseudo
 /codon_start=1
 /transl_table=11
 /db_xref="PID:e335008"
 /db_xref="PSEUDO:CAB11371.1"
 1478..2221
 /gene="MLCB22.07"
 1478..2221
 /gene="MLCB22.07"
 /note="MLCB22.07, unknown, len: 247 aa; highly similar in
 C-terminus to M. tuberculosis Y00H_MYCTU Q10395 very

JOURNAL	COMMENT
Nature 409 (6823), 1007-1011 (2001)	
MEDLINE	
21128732	
PUBMED	
11234002	
REFERENCE	
2 (bases 1 to 348450)	
AUTHORS	
TITLE	
JOURNAL	
Submitted (20-FEB-2001) Submitted on behalf of the Mycobacterium	
leprae sequencing teams, The Sanger Centre, Wellcome Trust Genome	
Campus, Hinxton, Cambridge, CB10 1SA, UK Unite de Genetique	
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,	
75724, Paris Cedex, France. E-mail: parkhill@sanger.ac.uk	
Notes:	
Details of M. leprae sequencing at the Sanger Centre are available	
from http://www.sanger.ac.uk/Projects/M_leprae/ A relational	
database containing the M. leprae sequences is available from	
http://genolist.pasteur.fr/Leprama/ .	
location/Qualifiers	
1. 348450	
organism="Mycobacterium leprae"	
/mol_type="genomic DNA"	
/strain="TN"	
/db_xref="taxon:1769"	
551..1786	
/gene="ML0842"	
551..1786	
/note="Similar to Mycobacterium tuberculosis hypothetical	
44.6 kDa protein Rv1464 or MV007.11 TR:053155	
(EMBL:AL021184) (417 aa) fasta scores: E(): 0, 43.6% id in	
408 aa and to many other bacterial nifs-homologues, e.g.	
Bacillus subtilis yurW protein yurW TR:032164	
(EMBL:D99120) (406 aa) fasta scores: E(): 0, 46.7% id in	
405 aa. Previously sequenced as TR:032975 (EMBL:Z98741).	
Contains Pfam match to entry PF00266 aminotran_5,	
Aminotransferases class-V.	
Similar to ML0117, ML0596 and ML1708"	
/codon_start=1	
/transl_table=11	
/product="conserved hypothetical protein"	
/protein_id="CAC31223.1"	
/db_xref="GI:13092923"	
/translation="MDVDPIQVADPFILOTYVNGKPLIMPNAATQPOVVDR	
LSFFVHENSNTIRAHAEHLARATDAVERRETRARRTGAKADEITFAKTTANIN	
VAAWGSKHLPQGEDEVITLLEHRANIVPQQLSGQVALKVAPOVDAGLLNSEEF	
DLAGPRTKLVAAQVNSALGTVIGEKIVELGHRGAVLIDGASIPHPINSEIGL	
ADPFVSGHKIVGPTGTVLGCEDVETEMPQGGNMIVDTLSLTVQGPNNKFEF	
AGCNITADAVGALRYRVERGVORFASHEOALLDVAATRIADIPVRLVGTATK	
SVSPFLAGHPELEVGKALNABGIIAVRAGHCAQPIVRLGLEATVPSPAFNTYEE	
IDVFINVRIALGAGTNG"	
653..1738	
/gene="ML0842"	
/note="Pfam match to entry PF00266 aminotran 5,	
Aminotransferases class-V, score 23.70, E-value 1.7e-08"	
1908..2395	
/gene="ML0843"	
/note="synonym: lppP"	
/pseudo	
1908..2395	
/gene="ML0843"	
/note="Possible pseudogene of M. tuberculosis orthologue	
lppP (Best blastx score 294)"	
/pseudo	
/codon_start=1	
/transl_table=11	
/product="lipoprotein (pseudogene)"	
2549..2555	
/note="possible RBS"	
2559..4112	
/gene="ML0844"	
/note="synonym: narK"	
2559..4112	
/gene="ML0844"	
/note="Similar to Mycobacterium tuberculosis putative	

nitrite extrusion protein nark1 or Rv2329c or MTCY312.05
 TR:P71883 (EMBL:Z79702) (515 aa) fasta scores: E(): 0,
 69.3% id in 488 aa and to Escherichia coli nitrite
 extrusion protein 2 naru SW:NAU_ECOLI (P37758; P77596)
 (462 aa) fasta scores: E(): 0, 36.6% id in 459 aa. There
 is a frameshift near the C-terminus relative to the M.
 tuberculosis homologue. Previously sequenced as TR:032974
 (EMBL:Z98741). Contains hydrophobic, probable
 membrane-spanning regions."
 /codon_start=1
 /transl_table=11
 /product="putative nitrite extrusion protein"
 /protein_id="CAC31225.1"
 /db_xref="GI:13092924"
 /translation="MDQVYLQNEESLYVGRSKTVPIIDHRSLKSPFRLRRAC
 RISHMDPEDQAAEAGNKTIARLNLMSVLTALHGSVNTLPMWELPMKQVPSA
 GKRFLATTATLVGACLRVPSLATLIFEGSRWALPSVWLLIPTATVVLAVHGLP
 LMPYLACALTLGSGNPAASMTNAAFPRLKGAFLGAGAGLVSIQVWTRDS
 VLVASVDRKPYIVCGLYVLLIAGIAMLFNNDIEHRIGVITRPVLFVVSTRDS
 NVIALIYLASFGSIFSPAGVLETNVPAGOSPAQALAAELAFPTLAAVAR
 FSGRLADRLGSGRVTLVVGAMVFAAGLGVLTIGSGRVCPIRGVMAVYFAGIT
 LFLGSGNGSVYKMTPTIPKASHSLGINDDECDMSRVISGVYGPVAGACGV
 GTDLARESYLNTGVTAAFWIPMLCTAAGVLTWMTVCFLPGLNDEANMAFAS
 VASRTRRG"
 /complement(4100..4681)
 /gene="ML0845"
 /complement(4100..4681)
 /gene="ML0845"
 /note="Unknown function. Similar to part of some acyl-CoA
 oxidases e.g. Arabidopsis thaliana acyl-CoA oxidase cx2
 TR:065201 (EMBL:AF057043) (692 aa) fasta scores: E():
 1.5e-06, 35.2% id in 125 aa. Previously sequenced as
 TR:032973 (EMBL:Z98741) (193 aa) fasta scores: E(): 0,
 99.5% id in 193 aa."
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="CAC31226.1"
 /db_xref="GI:13092925"
 /translation="MELRPTNAASERVTRTATENTLQRMVGARODESKKSLFNRGT
 OTIPKCKDCLSAVARRRQYKSKEMPAFPPTNTYQDHLAAEHTRVYLFETLA
 GIESCNDFARKLIDVCDLVASVARKGTSSIDTSPSALKALAKSGPMPMSAAL
 CRDVGRFSESSQSCVTPRCRTSRICPSLQPR"
 /complement(4959..5480)
 /gene="ML0846"
 /pseudo
 /complement(4959..5480)
 /gene="ML0846"
 /note="Possible pseudogene similar to M. tuberculosis
 paralogue Rv1747 (best blastx score 127)"
 /pseudo
 /codon_start=1
 /transl_table=11
 /product="ABC transporter (pseudogene)"
 /complement(5522..5848)
 /gene="ML0847"
 /note="Possible pseudogene of M. tuberculosis orthologue
 Rv2327 (best blastx score 184)"
 /pseudo
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein (pseudogene)"
 /complement(5906..8080)
 /gene="ML0848"
 /complement(5906..8080)
 /gene="ML0848"
 /note="Similar to Mycobacterium tuberculosis hypothetical
 ABC transporter ATP-binding protein Rv3326c or MTCY312.08
 SW:YN26 MYCTU (P71886) (697 aa) fasta scores: E(): 0,
 76.9% id in 697 aa. Shares similar domains with many

ABC-type transporters e.g. Streptomyces roseofulvus ATPase
 component of putative ABC transporter fird TR:068910
 (EMBL:AF053302) (524 aa) fasta scores: E(): 8.4e-14, 31.5%
 id in 495 aa and Synecchococcus sp. nitrate transport
 ATP-binding protein NtrD ntrd SW:NTRD_SYNP7 (P38046) (274
 aa) fasta scores: E(): 1.6e-12, 34.8% id in 198 aa.
 Previously sequenced as TR:032971 (EMBL:Z98741). Contains
 hydrophobic, possible membrane-spanning regions. Contains
 2 Pfam matches to entry PF00005 ABC tran, ABC transporter.
 Contains 2 x P800017 ATP/GTP-binding site motif A
 (P-loop). Contains 2 x P800211 ABC transporters family
 signature."
 /codon_start=1
 /transl_table=11
 /product="ABC transporter"
 /protein_id="CAC31229.1"
 /db_xref="GI:13092926"
 /translation="MDLCTHRMALLCOATRCRHWKTQGLGCAADLPATGTTGKV
 TADPRHLHRSGLQPELQASVLAACVTAIVSVVPPAALALGTPMELLA
 YRFRVMTMVAAGVAPLITGLGAVAVNSAYIGLGVGRKOGTLVLA
 LFAGLAFGAANVAVLVGRLHLFKAMTAVDGIATLTMELPMVAVOLKRYEAD
 GLQHPWMLIGYFVITLIVSLIGWVLSPLVERIRDIDVHKLDAPACNDAVGP
 VPMIDKTRFTRFPAQDALARVSDLDIAVGEHVAVTGANGSKTTIMILAEPTSG
 TVDRGAVGLGSLGTAVALQHPESQVLRVADVVGNGLEPGTDVDRRLREYGLD
 AFAERDTGSLSGELQRLAALAREBSLLIADVTSVMVRQGDALGLVSLTKR
 HPIALVHTHYNNEADRTINLSQSDSDNMGSETVAPVSTVAVDRPHAPVLELV
 GVGEHYGSGTPEAKALHDISFVNGSGGVLYGNGSGKSTIAMVIMGPTTGAC
 LIDGRPTHEHGAVALSFQARALGRVLEVDVNSAGFSRDRDRAAALGVGLD
 PALAKRIDQSGCMRRVYVLAIGLACSPRLITIDEPRLAGIDAVSGRGLRLLEDR
 EKLTVVTSIDHPGLBDCVPTVRLNGALVESVITAGTS"
 /gene="ML0848"
 /note="Pfam match to entry PF00005 ABC tran, ABC
 transporter, score 112.30, E-value 9e-30"

misc_feature

Query Match 68.7%; Score 20.6; DB 1; Length 348450;
 Best Local Similarity 85.2%; Pred. No. 1.8e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCTGGTAGTGTACGACTCCTGCGAG 28

Db 5076 CTTAAGTAGTGCAGACCTCCTGCGAG 5102

RESULT 3

AE003979

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 11282)

Xylella fastidiosa 945C

Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

Xanthomonadaceae; Xylella.

Alvaranga, R., Alves, L.M., Araya, J.E., Baia, G.S., Baptista, C.S.,

Barros, M.H., Bonaccorsi, E.D., Bordin, S., Bove, O.M., Brito, M.R.,

Bueno, M.R., Camargo, A.A., Camargo, L.E., Carraro, D.M., Carrier, H.,

Colauto, N.B., Colombo, C., Costa, F.F., Costa, M.C., Costa-Neto, C.M.,

Coutinho, L.L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorri, H.,

Faccin, A.P., Ferreira, A.J., Ferreira, V.C., Ferro, J.A.,

Fraga, J.S., Franca, S.C., Franco, M.C., Frohme, M., Furlan, L.R.,

Gardner, M., Goldman, G.H., Goldman, M.H., Gomes, S.L., Gruber, A.,

Ho, P.L., Hohnsbeil, J.D., Junqueira, M.L., Kemper, S.L., Kitajima, J.P.,

and Marino, C.L.

The genome sequence of the plant pathogen Xylella fastidiosa. The

Xylella fastidiosa Consortium of the Organization for Nucleotide
 Sequencing and Analysis
 JOURNAL MEDLINE
 PUBMED
 Nature 406 (6792), 151-157 (2000)
 20365717
 10910347


```

/gene="XF1498"
/complement(7506. . 9266)
/gene="XF1498"
/note="similar to SP|P52673 (percent identity: 52 %/query
alignment coverage: 93.0 %/subject alignment coverage:
97.5 %) : identified by sequence similarity; putative; ORF
located using Glimmer/Rasfinder"
/codon_start=1
/translation_table=1
/product="NADPH-sulfite reductase, iron-sulfur protein"
/protein_id="AAF84307.1"
/db_xref="GI:9106527"
/translation="MSPTSLSLHPGSPAPMSSSIEDIKGSKSHRSGSLLESAN
PTGLAHSDDTLIKRHSYQDDRDRLPESRRKQKALPAQPMIRTRTPEGVITPQOM
LQQLAIRVYNHSLRVTTRQAFQPHGVTKEELATQALNALITLTAAGSVNNV
OVANAPKISRAHADLYTDAALSHLLPNTAAVYIWLDRKQAGAGEEBRPYGPYH
LPKRKKIFSAAPRINDVDVFANDGLFATVNDTLGYNALIGGAGCTTGDDTPMR
VNIIGPTIRADLIALISPAIVTTRDGRNTRRKARPKYATIDRGLDCTVGRILQORA
GITLQAPRPVPEHNGDRYGVITEGEDRNMHTLSLPRGRADTRESSPLSGPAINQL
GITQGRMTPNQNVYAGISPGQRATIDLVITYQVLDGNGAPPLALAHNAACALPTC
GLAMAEARYIPDPFVNTKQPTLEKRYGLAEKRLILRISGCPNGCSRPPLAIALVGAP
GRVYNMLGGDGRGRNLTLVRENTTEETILLAAEPLILGRYQOKRLPGEFGDPLHRTG
IIALPPYPTHHVVISSTLOA"
/complement(9336. . 11174)
/gene="XF1499"
/complement(9336. . 11174)
/gene="XF1499"

```

	Query Match	68.0%	Score 20.4	DB 1	Length 11882;
	Best Local Similarity	80.0%	Pred. No. 1.7e+02;		
	Matches 24; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0;
Oy	1 CCCTGCGTAGTGTGCACACTCTCTGCAGGG	30			
b	7107 CCTGCGCATCGGTTGCACCTTGTCAGGG	7136			

RESULT 4			
AE012555			
LOCUS	302835 bp	DNA	linear BCT 29-JAN-2003
DEFINITION	<i>Xylella fastidiosa</i> Temeculal, section 3 of 9 of the complete		

VERSION	AE012555.1	GI:28056497
KEYWORDS		
SOURCE	Xylella fastidiosa Temeculal	
ORGANISM	Xylella fastidiosa Temeculal	

REFERENCE
1 (bases 1 to 302835)
Van Sluys, M.A., de Oliveira, M.C., Monteiro-Vitorello, C.B.,

TITLE Comparative Analyses of the Complete Genome Sequences of *Pierce's* Civerolo, E. L., Simpson, A. J. G., Almeida Jr., N. F., Setubal, J. C. and Kitajima, J. P.

JOURNAL
PUBMED
REFERENCE
AUTHORS

Zacarias
J. Bacteriol. 185 (3), 1018-1026 (2003)
12533478
2 (bases 1 to 302835)
Van Sluys, M.A., de Oliveira, M.C., Monteiro-Vitorello, C.B.,
Miyaki, C.Y., Furlan, L.R., Camargo, L.E.A., da Silva, A.C.R.,
Moon, D.H., Takita, M.A., Lemos, E.C.M., Machado, M.A., Ferro, M.I.T.,

TITLE	Direct Submission
JOURNAL	Submitted (17-DEC-2041) Instituto de Biotecnologias, Universidade de Sao Paulo, Rua do Matão, 277, Sao Paulo, SP 05508-900, Brazil
FEATURES	Location/Qualifiers
SOURCE	1. 302835

```
gene      359..4018
          /gene="p11y1"
          /locus_tag="PD0502"
CDS       359..4018
          /gene="p11y1"
```

gene	CDS
gene	complement (4427..4915) /locus_tag="pp0503" complement (4427..4915) /locus_tag="pp0503" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer" /codon_start=1 /transl_table=1 /product="conserved hypothetical protein" /protein_id="AA028178.1" /db_xref="GI:28056499" /translation="MLRPAIAGFFVYLLTPWKAISTHYSVNTFPNOSTPTVYTPIN TFNPQPPLPIPGKALSORNRELYNDPMLHLRVNSRQSIATKKTSTESTANSNN KTEIDYKQTIYDNTPMRFNNNOGKMTADEFAMMKARGVHVIPTQTNNNVQSVB AK" complement (5836..6225) /locus_tag="pp0504"
gene	

```
CDS      complement(5836..6225)
        /locus_tag="PD0504"
        /note="identified by sequence similarity; putative; ORF
        located using Blastx/Glimmer"
        /codon_start=1
        /transl_table=11
        /product="conserved hypothetical protein"
        /protein_id="AA028379.1"
        /db_xref="GI:28056500"
        /translation="MNKFTAMTRALAKHTMDLAQOANDRIRRTAPQASOMLKSQALE
        VYKSRKYTGKLIKBNPSIAVAVAVSGLAGYMYRKYKMHKKKIYINGGEMNL
        EHADVYRBSIHRHRSAAPTDNNSG"
        complement(6491..9001)
        /gene="hrpB"
        /locus_tag="PD0505"
        complement(6491..9001)
        /gene="hrpB"
        /locus_tag="PD0505"
        /note="identified by sequence similarity; putative; ORF
        located using Blastx/Glimmer"
        /codon_start=1
        /transl_table=11
        /product="ATP-dependent helicase"
        /protein_id="AA028380.1"
        /db_xref="GI:28056501"
        /translation="MCRMNHMLPIDALLPATTSLSNHPRLLEAPPAGACTTQVPL
        ALLDAPMLNGRSIILBPRRYAVRNALFMAROCGEAGLGYHIFRESKVSARTI
        HVTGGILITRIODDPLDGLGALLPDPFPHRLASDGLALADLVOTLPEDLRTIV
        MSATLQCGRLAQLAPRLTSEGSAVPTTTTFPRRPSLETQSRALIQHALDTHGQ
        DVLTLPGORETARLQALILEKTLSPELHVMPLHGLPQOETRVLEPDQGRRLVLA
        TNVASSITLPGISVIDSGLABEPAYDPTGFTLDDVASTIQASADQAGAGRLAP
        GMAYRSLPQSOHLEAORRELIQTELSGALALELWGSSTLFIIDPPSGALASAREL
        LQRLGALISNTSTLPLGKMLAIGHPRLAAMLAPHDPOGALACDLALLDPALEH
        RSGGALISMRALAEKSGCLPHDTKSTLKALIDATHOMERRIRCDATPDPAIEAH
        TLGNLTATPERRIARHSDDPLYVLANGRRMLDMSALHGEWITASELHETKD
        IQALCGPLDERKLDADPERRIQHDVLRWDRKRALVAQESCFDRIVLDRPAGRI
        DPVOANVLTAQVRLGLDLSWTEHRLQWQARAVALHWMIPBLPLSNTTLATL
        DTLWLPAPGKRLDALDEAFAEQAVKSLISTQHALIDRHVPAHIIVPSGGRHIDY
        GLADGTPQPPVLSYQBLFGLPDTPIAKRRIPLTLHLSPGGRPLQVTDLSFV
        TNYEYVKKEMKGRYKPKPMPDDPWTATATHTKSPGK"
        complement(9502..9990)
        /locus_tag="PD0506"
        complement(9502..9990)
        /locus_tag="PD0506"
        /note="identified by sequence similarity; putative; ORF
        located using Blastx/Glimmer"
        /codon_start=1
        /transl_table=11
        /product="conserved hypothetical protein"
        /protein_id="AA028381.1"
        /db_xref="GI:28056502"
        /translation="MLMLAPFLTNGMKRPOVQAQEAQAKLIHGETRIANANAL
        IPHSLPLGNWGFYTDGRBLVVGPGQIPACIRIPLDKGVCGAARQRTORITDID
        TPDHIACDASASRSELVPLKNDTLIGVLIDSPTCARFSDQDTGISLARVSN
        LR"
        10282..12135
        /locus_tag="PD0507"
        10282..12135
        /locus_tag="PD0507"
        /note="identified by sequence similarity; putative; ORF
        located using Blastx/Glimmer"
        /codon_start=1
        /transl_table=11
        /product="pathogenicity protein"
        /protein_id="AA028382.1"
        /db_xref="GI:28056503"
        /translation="MEVLRCFNMNVRFITLSILPRYTAQMLRFRIPMLSLWLFASLA
        QANATVDKIQIKGLDKDVEEMKNITVSLSYVSGKEGSRLEYLLSQAPQTRQ
        ALBPGYAPAIRIDAPRONDHITVIVVDKEPVRIRQAHVAMGAADHMYORDI
        EDFPKLGEIRNHPITYEASKYRITRBLAEGCFADPTFRQRYEYTRATYAADIDINE
        SGRRTDNGPRFHHIDYFHGQLFNPLVVBEGSYFHEGDLRKSULTKIDTFSSIDIO
        KPFEADSBGNVPDVKLERAKSLYTAGISYSSSSGALRAGVRRRYNMAAGHMNA
        RLDVYQNRKSLTAYOIPAIFKWLDMYIFSAVAYDEQTOYIDLBNVKLSAASGOINR
```

```
HLTATASIALNLRERMYAADGTNTVAYQHSITLVYPOLEASYVDVDAATPPRNGSAAT
VILRGASAFGSKSNFTQLHQLRMFHLGSSSRLILGEGVTTWASDLVAMPPLRF
FAGVNSIRGAFPEVGPRTXPDPAFALGAGVPSAGAEYVYKGGFPGGAVFVDSG
SAFNRPDHNTGIGIGLRYRSPVGPVRVDIARGLNKPKPSKIQLYIDIANL"
12414..15953
/locus_tag="PD0508"
12414..15953
/locus_tag="PD0508"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="pathogenicity protein"
/protein_id="AA028383.1"
/db_xref="GI:28056504"
/translation="MDPIMRLILRLIDVLEMSNASTLVSQSEPPKLPSPMDVLP
SINYPALQITITITIDNRIRSKQPLILHRRIGKAVNVAGSELEIQALVDSRGN
RADSYRPRHDYSDTLTRTAVSPSRGPPSPFGVIGDLDRLQALAGRAPVLAQ
DVNLGRSDPRTSPASAEHPFGLLPGELTGPPTAPVLAENVHAAQOGKADLKGQ
FKQGEOTIVLQPSRLQLANQMLNWEPLVIDAFGGRTNLRGVDIRDPANTSPHPSVNA
```

Query Match 68.0%; Score 20.4; DB 1; Length 302835;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCTGCGTAGTGTACGACCTCTGCAGG 30
||||| ||||| ||||| ||||| |||||
Db 276037 CCTGCGCATGCGTTCGACCTTGTGCAGG 276066

RESULT 5
LMFLCHR34_04
WCOMMENT"
Sequence split into 18 fragments

Fragment Name	Begin	End	Accession
LMFLCHR34_00	1	110000	LMFLCHR34_00
LMFLCHR34_01	100001	210000	LMFLCHR34_01
LMFLCHR34_02	200001	310000	LMFLCHR34_02
LMFLCHR34_03	300001	410000	LMFLCHR34_03
LMFLCHR34_04	400001	510000	LMFLCHR34_04
LMFLCHR34_05	500001	610000	LMFLCHR34_05
LMFLCHR34_06	600001	710000	LMFLCHR34_06
LMFLCHR34_07	700001	810000	LMFLCHR34_07
LMFLCHR34_08	800001	910000	LMFLCHR34_08
LMFLCHR34_09	900001	1010000	LMFLCHR34_09
LMFLCHR34_10	1000001	1110000	LMFLCHR34_10
LMFLCHR34_11	1100001	1210000	LMFLCHR34_11
LMFLCHR34_12	1200001	1310000	LMFLCHR34_12
LMFLCHR34_13	1300001	1410000	LMFLCHR34_13
LMFLCHR34_14	1400001	1510000	LMFLCHR34_14
LMFLCHR34_15	1500001	1610000	LMFLCHR34_15
LMFLCHR34_16	1600001	1710000	LMFLCHR34_16
LMFLCHR34_17	1700001	1720777	LMFLCHR34_17

Continuation (5 of 18) of LMFLCHR34 from base 400001 (AL499623 Leishmania major chromosome)

Query Match 67.3%; Score 20.2; DB 2; Length 110000;
Best Local Similarity 88.0%; Pred. No. 2.5e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CCTGCGTAGTGTACGACCTCTGC 26
||||| ||||| ||||| ||||| |||||
Db 32750 CCTGCGTGTGGAGACGACCTCTGC 32774

RESULT 6
AC074063/c
AC074063/c
LOCUS AC074063 68974 bp DNA linear HTG 13-JUL-2000
DEFINITION Mus musculus chromosome 18 clone RP22-32123 map 18, LOW-PASS
SEQUENCE SAMPLING.
ACCESSION AC074063.1 GI:9154848
VERSION AC074063.1
KEYWORDS HTG; HTGS; PHASE0.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Birren, B., Linton, L., Nusbaum, C., and Lander, E.

JOURNAL Mus musculus chromosome 18, clone RP2-32123

REFERENCE Unpublished

AUTHORS 2 (bases 1 to 68974)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bachtien, V., Bede, F., Boguslavsky, L., Bouckgalter, B., Brown, A., Burkett, G., Campatano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J. S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-pierre, N., Grant, G., Hages, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPherson, R., Melgrim, J., Menues, L., Mihova, T., Miranda, C., Menga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Plesni, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Teste, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (13-JUL-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smit, A. P. A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submission@genome.wi.mit.edu

Project Information

Center project name: L5920

Center clone name: 32_I_23

NOTE: This record contains 81 individual

* sequencing reads that have not been assembled into

* contigs. Runs of N are used to separate the reads

* and the order in which they appear is completely

* arbitrary. Low-pass sequence sampling is useful for

* identifying clones that may be gene-rich and allows

* overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone

* will be sequenced to completion. In the event that

* the record is updated, the accession number will

* be preserved.

1 740: contig of 740 bp in length

* 741 840: gap of 100 bp

* 841 1602: contig of 762 bp in length

* 1603 1702: gap of 100 bp

* 1703 2456: contig of 754 bp in length

* 2457 2556: gap of 100 bp

* 2557 3294: contig of 738 bp in length

* 3295 3394: gap of 100 bp

* 3395 4122: contig of 728 bp in length

* 4123 4222: gap of 100 bp

* 4223 4986: contig of 764 bp in length

* 4987 5086: gap of 100 bp

* 5087 5826: contig of 740 bp in length

* 5827 5926: gap of 100 bp

* 5927 6678: contig of 752 bp in length

* 6679 6778: gap of 100 bp

* 6779 7508: contig of 730 bp in length

* 7509 7608: gap of 100 bp

7609 8348: contig of 740 bp in length

* 8349 8448: gap of 100 bp

* 8449 9191: contig of 743 bp in length

* 9192 9291: gap of 100 bp

* 9292 10044: contig of 753 bp in length

* 10045 10144: gap of 100 bp

* 10145 10890: contig of 746 bp in length

* 10891 10990: gap of 100 bp

* 10991 11746: contig of 756 bp in length

* 11747 11846: gap of 100 bp

* 11847 12602: contig of 756 bp in length

* 12603 12702: gap of 100 bp

* 12703 13452: contig of 750 bp in length

* 13453 13552: gap of 100 bp

* 13553 14300: contig of 748 bp in length

* 14301 14400: gap of 100 bp

* 14401 15148: contig of 748 bp in length

* 15149 15248: gap of 100 bp

* 15249 16021: contig of 773 bp in length

* 16022 16121: gap of 100 bp

* 16122 16852: contig of 731 bp in length

* 16853 16952: gap of 100 bp

* 16953 17711: contig of 753 bp in length

* 17712 17811: gap of 100 bp

* 17812 18566: contig of 755 bp in length

* 18567 19417: contig of 751 bp in length

* 19418 19517: gap of 100 bp

* 19518 20283: contig of 766 bp in length

* 20284 20383: gap of 100 bp

* 20384 21132: contig of 749 bp in length

* 21133 21322: gap of 100 bp

* 21323 21988: contig of 756 bp in length

* 21989 22088: gap of 100 bp

* 22089 22858: contig of 770 bp in length

* 22859 22958: gap of 100 bp

* 22959 23703: contig of 745 bp in length

* 23704 23803: gap of 100 bp

* 23804 24557: contig of 754 bp in length

* 24558 24657: gap of 100 bp

* 24658 25410: contig of 753 bp in length

* 25411 25510: gap of 100 bp

* 25511 26274: contig of 764 bp in length

* 26275 26374: gap of 100 bp

* 26375 27147: contig of 773 bp in length

* 27148 27247: gap of 100 bp

* 27248 28007: contig of 760 bp in length

* 28008 28107: gap of 100 bp

* 28108 28862: contig of 755 bp in length

* 28863 28962: gap of 100 bp

* 28963 29709: contig of 747 bp in length

* 29710 29809: gap of 100 bp

* 29810 30568: contig of 755 bp in length

* 30569 30668: gap of 100 bp

* 30669 31493: contig of 732 bp in length

* 31494 31993: gap of 100 bp

* 31994 32251: contig of 758 bp in length

* 32252 32351: gap of 100 bp

* 32352 33105: contig of 754 bp in length

* 33106 33205: gap of 100 bp

* 33206 33966: contig of 761 bp in length

* 33967 34066: gap of 100 bp

* 34067 34816: contig of 750 bp in length

* 34817 34916: gap of 100 bp

* 34917 35662: contig of 746 bp in length

* 35663 35762: gap of 100 bp

* 35763 36534: contig of 772 bp in length

* 36535 36634: gap of 100 bp

* 36635 37399: contig of 765 bp in length

* 37400 37499: gap of 100 bp

* 37500 38246: contig of 747 bp in length

* 38247 38346: gap of 100 bp

* 38347 39098: contig of 752 bp in length


```

* 39099 39198: gap of 100 bp
* 39199 39937: contig of 739 bp in length
* 39938 40037: gap of 100 bp
* 40038 40783: contig of 746 bp in length
* 40784 40883: gap of 100 bp
* 40884 41572: contig of 789 bp in length
* 41573 41772: gap of 100 bp
* 41773 42535: contig of 763 bp in length
* 42536 42635: gap of 100 bp
* 42636 43403: contig of 768 bp in length
* 43404 43503: gap of 100 bp
* 43504 44255: contig of 752 bp in length
* 44256 44355: gap of 100 bp
* 44356 45086: contig of 731 bp in length
* 45087 45186: gap of 100 bp
* 45187 45916: contig of 730 bp in length
* 45917 46016: gap of 100 bp
* 46017 46741: contig of 725 bp in length
* 46742 46841: gap of 100 bp
* 46842 47597: contig of 756 bp in length
* 47598 47698: gap of 100 bp
* 47699 48450: contig of 753 bp in length
* 48451 48550: gap of 100 bp
* 48551 49303: contig of 753 bp in length
* 49304 49403: gap of 100 bp
* 49404 50159: contig of 756 bp in length
* 50160 50259: gap of 100 bp
* 50260 51011: contig of 752 bp in length
* 51012 51111: gap of 100 bp
* 51112 51860: contig of 749 bp in length
* 51861 51960: gap of 100 bp
* 51961 52723: contig of 763 bp in length
* 52724 53589: contig of 766 bp in length
* 53590 53689: gap of 100 bp
* 53690 54439: contig of 750 bp in length
* 54440 54539: gap of 100 bp
* 54540 55302: contig of 763 bp in length
* 55303 55402: gap of 100 bp
* 55403 56157: contig of 755 bp in length
* 56158 56257: gap of 100 bp
* 56258 57008: contig of 751 bp in length
* 57009 57108: gap of 100 bp
* 57109 57880: contig of 772 bp in length
* 57881 57980: gap of 100 bp
* 57981 58742: contig of 762 bp in length
* 58743 58842: gap of 100 bp

```

```

Query Match      64.7%; Score 19.4; DB 2; Length 68974;
Best Local Similarity 79.3%; Pred. No. 5.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

Cy 1 CCTCGCGTAGTGTGACGACCTCTCGAGG 29
Db 30913 CTTGTGAGGTGTGCGAGACCTCTCGAGG 30885

```

```

RESULT 7
AC002315/c 110079 bp DNA linear ROD 20-FEB-1998
LOCUS Mouse BAC-146N21 Chromosome X contains iduronate-2-sulfatase gene;
DEFINITION complete sequence.
ACCESSION AC002315
VERSION AC002315.1 GI:2258164
KEYWORDS
ORGANISM Mus musculus (house mouse)
SOURCE Mus musculus
MUTAGENESIS
REFERENCE Muzny,D., Anzari-Lari,M.A., Timms,K.M., Yu,W., Dugan,S., Lu,J.,
AUTHORS Shen,Y., Kowland,K., Liu,W., Perez,L., Ding,Y., Haywood,M.,
Jain,A., Leal,B., Logan,O., Nguyen,V., Savage,L., Shen,H.,

```

```

TITLE
JOURNAL Direct Submission
REFERENCE 2 (bases 1 to 110079)
AUTHORS Timms,K.M., Hackett,L., Edwards,F.J., Lu,J., Muzny,D.M., Miller,W.,
and Gibbs,R.A.
TITLE Cross-species sequence comparison of the IDS region
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 110079)
AUTHORS Chiu,M.W.
JOURNAL Direct Submission
COMMENT Submitted (15-JUN-1997) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
Sequencing is completed to a minimum standard of double strand
coverage with a minimum of 2 clones and 2 reads with no ambiguities
or 2 chemistries with a minimum of 2 clones and 3 reads with no
ambiguities. If the sequence quality does not meet this standard,
it will be indicated in the annotation.
The repeat regions shown were identified using RepeatMasker by
Adrian Smit.
Sequence similarities were identified using Powerblast by Jinghui
Zhang.
Exon/Intron boundaries of identified genes were chosen if there
were canonical splice junctions that maintained sequence continuity
across the splice junctions.

```

FEATURES

source

1..110079

Location/Qualifiers

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="X"
/clone="BAC-146N21"
244..561
/rpt_family="LTR/MaLR"
repeat_region
567..2305
/rpt_family="LTR/MaLR"
repeat_region
2306..2646
/rpt_family="LTR/MaLR"
repeat_region
2647..3101
/rpt_family="LTR/Retroviral"
repeat_region
3102..3847
/rpt_family="Simple_repeat"
repeat_region
3848..5064
/rpt_family="Simple_repeat"
repeat_region
5065..5184
/rpt_family="LINE/L1"
repeat_region
5185..5457
/rpt_family="SINE/B2"
repeat_region
5458..6953
/rpt_family="SINE/B2"
repeat_region
6954..8199
/rpt_family="Simple_repeat"
repeat_region
8200..10861
/rpt_family="SINE/Alu"
repeat_region
10862..11185
/rpt_family="LTR/MaLR"
repeat_region
11186..11366
/rpt_family="SINE/B4"
repeat_region
11367..11718
/rpt_family="LTR/MaLR"
repeat_region
11719..11971
/rpt_family="Simple_repeat"
repeat_region
11972..12171
/rpt_family="SINE/B4"
repeat_region
12172..12276
/rpt_family="Low_complexity"
repeat_region
12277..12396
/rpt_family="SINE/B2"

```

```

repeat_region      14061..14118
                    /rpt_family="Simple_repeat"
unsubre            14096..14220
                    /note="Single strand coverage only"
repeat_region      14179..14201
                    /rpt_family="Low_complexity"
gene               join(14201..14287,15145..15281,15929..16106,17952..18040,
                    18934..19134,20646..20816,25808..25934,27982..28155)
                    /gene="TDS"
repeat_region      complement(14424..14541)
                    /rpt_family="Simple_repeat"
repeat_region      complement(14783..14914)
                    /rpt_family="Simple_repeat"
repeat_region      17009..17084
                    /rpt_family="SINE/MIR"
                    complement(17555..17697)
repeat_region      /rpt_family="SINE/Alu"
                    complement(17620..17719)
repeat_region      /rpt_family="LTR/MaLR"
                    complement(17821..17869)
repeat_region      /rpt_family="SINE/MIR"
                    complement(18187..18314)
repeat_region      /rpt_family="SINE/Alu"
                    18371..18499
repeat_region      /rpt_family="SINE/B4"
                    19306..19411
repeat_region      /rpt_family="SINE/Alu"
                    complement(19902..20388)
repeat_region      /rpt_family="Other/MER21_gro"
                    complement(21582..21611)
repeat_region      /rpt_family="Simple_repeat"
                    complement(21662..21761)
repeat_region      /rpt_family="sacRNA"
                    complement(21684..21761)
repeat_region      /rpt_family="SINE/ID"
                    21812..21849
repeat_region      /rpt_family="DNA/MER1_type"
                    complement(23849..23940)
repeat_region      /rpt_family="SINE/Alu"
                    25152..25220
repeat_region      /rpt_family="LTR/MaLR"
                    complement(26165..26376)
repeat_region      /rpt_family="SINE/B2"
                    complement(26379..26577)
repeat_region      /rpt_family="SINE/B2"
                    26995..27125
repeat_region      /rpt_family="SINE/B2"
                    27322..27505
repeat_region      /rpt_family="SINE/B2"
                    complement(27695..27734)
repeat_region      /rpt_family="Low_complexity"
                    29292..30131
repeat_region      /rpt_family="LINE/L1"
                    30204..30233
repeat_region      /rpt_family="Simple_repeat"
                    complement(31179..31235)
repeat_region      /rpt_family="Simple_repeat"
                    31236..31266
repeat_region      /rpt_family="Simple_repeat"
                    31579..31703
repeat_region      /rpt_family="Simple_repeat"
                    31705..32083
repeat_region      /rpt_family="LINE/L1"
                    33920..33985
repeat_region      /rpt_family="Low_complexity"
                    34396..34534
repeat_region      /rpt_family="Simple_repeat"
                    complement(34591..34664)
repeat_region      /rpt_family="DNA/MER1_type"
                    complement(34824..34916)
repeat_region      /rpt_family="Simple_repeat"
                    complement(35493..35633)
repeat_region      /rpt_family="LINE/L1"

```

```

repeat_region      complement(36159..36380)
                    /rpt_family="SINE/B2"
repeat_region      complement(36381..36489)
                    /rpt_family="LTR/MaLR"
repeat_region      36490..36804
                    /rpt_family="LTR/MaLR"
repeat_region      36805..38233
                    /rpt_family="LTR/MaLR"
repeat_region      38236..38585
                    /rpt_family="LTR/MaLR"
repeat_region      complement(38986..39048)
                    /rpt_family="SINE/Alu"
repeat_region      39226..39295
                    /rpt_family="Low_complexity"
                    complement(39309..39438)
repeat_region      /rpt_family="SINE/Alu"
                    complement(39439..39929)
repeat_region      /rpt_family="LINE/L1"
                    complement(40689..40779)
repeat_region      /rpt_family="SINE/Alu"
                    complement(41671..42216)
repeat_region      /rpt_family="LTR/MaLR"
                    complement(42325..42428)
repeat_region      /rpt_family="Simple_repeat"
                    42539..42710
repeat_region      /rpt_family="Simple_repeat"
                    42896..42980
repeat_region      /rpt_family="DNA/MER1_type"
                    complement(43056..43106)
repeat_region      /rpt_family="Low_complexity"
                    43286..43340
repeat_region      /rpt_family="Low_complexity"
                    43406..43513
repeat_region      /rpt_family="SINE/Alu"
                    43947..44060
repeat_region      /rpt_family="Simple_repeat"
                    46601..46704
repeat_region      /rpt_family="DNA/MER1_type"
                    46977..47046
repeat_region      /rpt_family="Simple_repeat"

Query Match      64.7%; Score 19.4; DB 10; Length 110079;
Best Local Similarity 79.3%; Pred. No. 5.9e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1  CCTGCTAGTGTGACCTCTGCGAG 29
DB      4399  CACTGCTAGTGTGACCTCTGCGAG 4371

RESULT 8
AC121312      163132 bp      DNA      linear      HTG 17-MAR-2003
LOCUS      Mus musculus clone RP24-34305, WORKING DRAFT SEQUENCE, 10 unordered
DEFINITION      pieces.
ACCESSION      AC121312.3 GI:28975922
VERSION      HTG: HTGS PHASE1: HTGS_DRAFT.
KEYWORDS      Mus musculus (house mouse)
SOURCE      Mus musculus
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 163132)
AUTHORS      Birren,B., Nusbaum,C. and Lander,E.
TITL      Mus musculus, clone RP24-34305
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 163132)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Bouhassira,B., Brown,A., Camarata,J., Campiano,A., Chang,J.,
Charato,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeCarliano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galegan,J., Gardyna,S.,

```

TITLE
JOURNAL
REFERENCE
AUTHORS

Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Kagot, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Pollard, V.,
Raymond, C., Retter, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strasas, N., Sudramanlian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (16-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 163132)

TITLE
JOURNAL
COMMENT

Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collamore, A., Cook, A., Cooke, P., Corum, B., DeArallano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Gajagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mahbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, J., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N.,
Rachupka, A., Ramaesmy, U., Raymond, C., Retter, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbins, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (17-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 17, 2003 this sequence version replaced g1:25956378.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Genome Center
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L24754
Center clone name: 343 J 5

Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 160194 bases at least Q40
Consensus quality: 161479 bases at least Q30
Consensus quality: 161963 bases at least Q20
Insert size: 16800; agarose-fp
Insert size: 162232; sum-of-contigs
Quality coverage: 8.4 in Q20 bases; agarose-fp
Quality coverage: 8.7 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.
1 872: contig of 872 bp in length
* 873 972: gap of 100 bp
* 973 2003: contig of 1031 bp in length
* 2004 2103: gap of 100 bp
* 2104 4258: contig of 2155 bp in length
* 4259 4359: gap of 100 bp
* 4359 6590: contig of 2232 bp in length
* 6591 14225: gap of 100 bp
* 14225 14225: contig of 7535 bp in length
* 14226 14325: gap of 100 bp
* 14326 28588: contig of 14263 bp in length
* 28589 28688: gap of 100 bp
* 28689 44413: contig of 15725 bp in length
* 44414 44513: gap of 100 bp
* 44513 82435: contig of 37922 bp in length
* 82436 82535: gap of 100 bp
* 82536 127168: contig of 44633 bp in length
* 127169 127269: gap of 100 bp
* 127269 163132: contig of 35864 bp in length.
Location/Qualifiers
1. 163132
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-34355"
/clone_1b="RP24-34355"
1. 872
/note="assembly_fragment"
/note="assembly_fragment"
973. 2003
/note="assembly_fragment"
2104. 4258
/note="assembly_fragment"
4359. 6590
/note="assembly_fragment"
6591. 14225
/note="assembly_fragment"
14326. 28588
/note="assembly_fragment"
28689. 44413
/note="assembly_fragment"
44514. 82435
/note="assembly_fragment"
82536. 127168
/note="assembly_fragment"
127269. 163132
/note="assembly_fragment"
clone_end:19
vector_side:right

BASE COUNT 50534 a 33420 c 33393 g 44882 t 903 others
ORIGIN

Query Match 64.7%; Score 19.4; DB 2; Length 163132;
Best Local Similarity 79.3%; Pred. No. 6.1e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OR 2 CCGCGTAGTGTGATGACCTCTCGACGG 30
Db 135188 CCGCGCCAGCGGTGACCTCTCGACGG 135216

RESULT 9
AC117138/c 176222 bp DNA linear HTG 19-NOV-2002
LOCUS Rattus norvegicus clone CH230-385L3, WORKING DRAFT SEQUENCE, 4
DEFINITION uncloned pieces.
AC117138
AC117138 5 GI:25073145
KEYWORDS HTGS_PHASE1; HTGS_DRAFT; HTGS_FUTITOR.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 176222)

Munry, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Anyalbech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Blawie, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Faller, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howell, S., Hulik, S., Hume, J., Idledit, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Kapachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kovacs, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzenhwa, L., Louised, H., Lorzado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmood, M., Mallory, K., Mangun, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiley, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Naik, L., Narkewicz, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokoleme, O., Okwona, G., Olarnpungson, A., Pal, S., Parks, K., Patel, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, B., Pu, L. L., Puzo, B., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivas, C., Rodkey, T., Rojce, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, M., Saverly, G., Scherer, S., Scott, C., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C., D. Smajd, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, A., Steimle, M., Strong, R., Sutton, A., Swatek, A., Tabor, P., Taylor, C., Taylor, R., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Valae, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wood, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 176222)

Worley, K. C.

Direct Submission

Submitted (08-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 176222)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GUCH

Center clone name: CH230-385L3

Summary Statistics

Assembly program: Phrap, version 0.990329

Consensus quality: 163801 bases at least Q40

Consensus quality: 166225 bases at least Q30

Consensus quality: 167951 bases at least Q20

Estimated insert size: 166810; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 4 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 3216: contig of 3216 bp in length

3317 14915: contig of 1159 bp in length

14916 15015: gap of unknown length

15016 139403: contig of 12438 bp in length

139404 139503: gap of unknown length

139504 176222: contig of 36719 bp in length.

Location/Qualifiers

1. 176222

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-385L3"

1. 1160

/note="wgs_end_extension"

clone_end:T7

3317. 4401

/note="wgs_end_extension"

clone_end:T7

10164. 11253

/note="wgs_end_extension"

clone_end:T7

13591. 14915

/note="wgs_end_extension"

clone_end:T7

15016. 16802

/note="wgs_end_extension"

clone_end:T7

17106. 18170

/note="wgs_end_extension"

clone_end:T7

33714. 34597

/note="clone_boundary"

clone_end:T7

site:

end_sequence: B2134038"

BASE COUNT 50120 a 34724 c 34881 g 49598 t 6799 others

ORIGIN

Query Match 64.7% Score 19.4; DB 2; Length 176222;

Best Local Similarity 79.3%; Pred. No. 6.1e+02;

[illegible]

```

repeat_region 28280..28490
                /rpt_family="L1"
repeat_region 28491..33172
                /rpt_family="L1_MM"
misc_feature 33178..33672
              /note="Region: Similar to U29402|MMU29402 Mus musculus
              acidic ribosomal phosphoprotein P1 mRNA, complete cds"
repeat_region 33682..35079
                /rpt_family="L1_MM"
repeat_region 37207..37292
                /rpt_family="RSINE1"
repeat_region 37310..37338
                /rpt_family="GA)n"
repeat_region 37354..37474
                /rpt_family="CATR)n"
repeat_region 37574..37726
                /rpt_family="RSINE1"
repeat_region 38061..38179
                /rpt_family="(TA)n"
repeat_region 39320..39492
                /rpt_family="B4"
repeat_region 39603..39661
                /rpt_family="(CA)n"
repeat_region 40026..40452
                /rpt_family="Lx2B"
repeat_region 40580..40624
                /rpt_family="(CA)n"
repeat_region 40951..41040
                /rpt_family="MT2B"
repeat_region 41143..41538
                /rpt_family="MTA"
repeat_region 42185..42262
                /rpt_family="(CAG)n"
repeat_region 43308..43316
                /rpt_family="(CAG)n"
repeat_region 43486..43974
                /rpt_family="L1"
repeat_region 43954..44894
                /rpt_family="L1_MM"
repeat_region 45112..45198
                /rpt_family="(CAA)n"
repeat_region 45459..45479
                /rpt_family="AT_rich"
repeat_region 46721..46836
                /rpt_family="(GAA)n"
repeat_region 46914..47150
                /rpt_family="MTD"
repeat_region 47213..47296
                /rpt_family="PB1D10"
repeat_region 50241..50324
                /rpt_family="MTC"
repeat_region 50328..51555
                /rpt_family="L1_MM"
repeat_region 51568..52378
                /rpt_family="L1_MM"
repeat_region 52395..52434
                /rpt_family="POLY_A"
repeat_region 52446..52503
                /rpt_family="L1_MM"
repeat_region 52520..52838
                /rpt_family="MTC"
repeat_region 53200..53245
                /rpt_family="(CAAAA)n"

```

```

Query Match 64.7% Score 19.4; DB 10; Length 194985;
Best Local Similarity 79.3%; Pred. No. 6.2e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

OY 1 CCTGCGTAGTGTACGACCTCTGCAGG 29
DB 117060 CACTGCTTAGTGCGCATGTCACTCTGCAGG 117088

```

```

RESULT 11
AL672026 200548 bp DNA linear ROD 19-SEP-2002
LOCUS Mouse DNA sequence from clone RP23-403011 on chromosome X, complete
DEFINITION sequence.
ACCESSION AL672026
VERSION AL672026.9 GI:22204360
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITL (bases 1 to 200548)
JOURNAL Direct Submission
Submitted (19-SEP-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone request: clonerequest@sanger.ac.uk
On Aug 11, 2002 this sequence version replaced gi:21955546.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
-----

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-403011 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBac3.5.

```

FEATURES
    Location/Qualifiers
        source
            1..200548
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /chromosome="X"
                /clone="RP23-403011"
                /clone_id="RPCI-23"

```

```

BASE COUNT 56146 a 44414 c 44985 g 55003 t
ORIGIN

```

```

Query Match 64.7% Score 19.4; DB 10; Length 200548;
Best Local Similarity 79.3%; Pred. No. 6.2e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

OY 1 CCTGCGTAGTGTACGACCTCTGCAGG 29
DB 10079 CACTGCTTAGTGCGCATGTCACTCTGCAGG 10107

```

```

RESULT 12
AC126569 217622 bp DNA linear HTG 10-MAY-2003
LOCUS Rattus norvegicus clone CH230-27C5, WORKING DRAFT SEQUENCE, 2
DEFINITION unordered pieces.
ACCESSION AC126569

```


ACCESSION AC113777
 VERSION AC113777.5 GI:25072715
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLOP.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 222869)
 Murny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooke, S., Amin, A., Angiano, D., Anvatsich, V., Aoyagi, A., Ayodeji, N., Baca, E., Baden, H., Baldwin, D., Bandaru, D., Bandaru, D., Barnes, M., Barnstead, M., Benahmed, F., Blevins, K., Blair, J., Blankenship, K., Blyth, P., Brown, M., Bryan, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dedrich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, J., Escott, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., George, G., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunatane, P., Haland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howell, S., Hulyk, S., Hume, J., Idelchik, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Kapachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovac, C., Kovacs, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, B., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Natr, L., Narkervic, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nuckeltemah, O., Okwunnu, G., Olamunagbon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L. L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Reiser, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajic, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wiczyska, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 222869)
 AUTHOR Morley, K. C.
 JOURNAL Direct Submission
 Submitted (05-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 222869)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 COMMENT On Nov 19, 2002 this sequence version replaced gi:2326985.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence scaffolds are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GSGF
 Center clone name: CH230-92115
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 217955 bases at least Q40
 Consensus quality: 218938 bases at least Q30
 Consensus quality: 219754 bases at least Q20
 Estimated insert size: 223449; sum-of-contigs estimation
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 113340: contig of 113340 bp in length
 * 113341 113440: gap of unknown length
 * 113441 203688: contig of 90248 bp in length
 * 203689 203788: gap of unknown length
 * 203789 222869: contig of 19081 bp in length.

FEATURES
 source
 1. 222869
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-92115"
 /note="wgs contig"

misc_feature
 113441..115233
 59606 a 48726 c 48850 g 63215 t 2472 others

BASE COUNT
 59606 a 48726 c 48850 g 63215 t 2472 others

ORIGIN
 Query Match 64.7% Score 19.4; DB 2; Length 222869;
 Best Local Similarity 79.3%; Pred. No. 6.2e+02;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Seq 1 CCCTGCTAGTGTACGACCTCTGCGAG 29
 Db 47915 CTCTGAGTTCTGGAGGAGCCTCTGCGAG 47943

RESULT 14
 EX294168/c 226382 bp DNA linear HTG 02-JUN-2003
 LOCUS Mus musculus chromosome X clone RP23-29M4, *** SEQUENCING IN
 DEFINITION PROGRESSES ***, 5 unordered pieces.
 ACCESSION EX294168
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLOP.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 17:02:34 ; Search time 150.297 Seconds
(without alignments)
538.821 Million cell updates/sec

Title: US-09-594-065-3
Perfect score: 30
Sequence: 1 cccctgcgtactgtacgacctccgcaggg 30

Scoring table:
IDENTITY NUC
Gapop 10-0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18.8	62.7	576	22	AAH67112
C 2	18.8	62.7	1368	18	AAH66463
C 3	18.8	62.7	2936	18	AAH66461
C 4	18.8	62.7	4454	18	AAH66462
C 5	18.8	62.7	349980	22	AAH68530
C 6	18.4	61.3	294	24	ABN17942
C 7	18.4	61.3	3653	24	AB199918
C 8	18.4	61.3	4705	24	ABA96146

9	18.4	61.3	4786	24	ABA96147
C 10	18.4	61.3	1503841	24	ABF00010
C 11	18.4	61.3	1503841	24	ABF001503
C 12	18.4	61.3	1503900	22	AAK95240
C 13	18.4	61.3	1503900	22	AAK95240
C 14	17.8	59.3	288	25	ABX82153
C 15	17.8	59.3	1035	24	ABL72764
C 16	17.8	59.3	1035	22	AAH66352
C 17	17.8	59.3	1766	21	AAH66299
C 18	17.8	59.3	2035	15	AAH66537
C 19	17.8	59.3	2035	24	ABE62374
C 20	17.8	59.3	3444	23	ABL29461
C 21	17.8	59.3	9205	23	ABL29461
C 22	17.6	58.7	2553	23	ABL10891
C 23	17.6	58.7	4829	25	ABX08797
C 24	17.6	58.7	11431	23	ABL19365
C 25	17.6	58.7	21837	22	AAH85996
C 26	17.4	58.0	243	23	AAH85688
C 27	17.4	58.0	262	22	AAH85857
C 28	17.4	58.0	262	22	AAH85857
C 29	17.4	58.0	262	22	AAH8458
C 30	17.4	58.0	262	22	AAH8458
C 31	17.4	58.0	325	21	AAH10137
C 32	17.4	58.0	799	22	AAH51449
C 33	17.4	58.0	799	24	ABO66773
C 34	17.4	58.0	823	22	AAH08485
C 35	17.4	58.0	823	25	ABH22209
C 36	17.4	58.0	878	24	ABH53342
C 37	17.4	58.0	1062	24	ABL90201
C 38	17.4	58.0	1196	23	AAH85689
C 39	17.4	58.0	1446	21	AAH52649
C 40	17.4	58.0	1478	22	AAH529780
C 41	17.4	58.0	1487	22	AAH60357
C 42	17.4	58.0	1508	22	AAH529805
C 43	17.4	58.0	1520	22	AAH58571
C 44	17.4	58.0	1569	24	AAH59892
C 45	17.4	58.0	1592	23	AAH85690

ALIGNMENTS

RESULT 1	AAH67112/c
ID	AAH67112 standard; DNA; 576 BP.
AC	AAH67112;
XX	26-SEP-2001 (first entry)
DT	
XX	C glutamicum coding sequence fragment SEQ ID NO: 2147.
XX	
XX	Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW	organic acid synthesis; ds.
XX	
OS	Corynebacterium glutamicum.
XX	
XX	EP1108790-A2.
XX	20-JUN-2001.
XX	
XX	18-DEC-2000; 2000EP-0127688.
XX	
XX	16-DEC-1999; 99UP-0377484.
XX	07-APR-2000; 2000UP-0159162.
XX	03-AUG-2000; 2000UP-0280988.
XX	
XX	(KEYW) KYOWA HAKKO KOGYO KK.
XX	
XX	Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX	Tateishi N, Senoh A, Ikeda W, Ozaki A;
XX	WPI; 2001-376931/40.

Modified human/mou
Human neurogulin 1
Human neurogulin 1
Human neurogulin-1
Human neurogulin-1
Corn ear-derived p
Corn tassal-derivate
P. putida oxygens
Human secreted pro
Sequence encoding
Colon adenocarcino
Drosophila melanog
Drosophila melanog
Angiogenesis-assoc
Drosophila melanog
Human immune/haema
DNA encoding novel
Human immune/haema
Human immune/haema
Human immune/haema
Human immune/haema
Human secreted pro
Human cDNA encodin
Human polynucleoti
Human cDNA clone (Breast cancer mark
Corn event MON180
Human polynucleoti
DNA encoding novel
Eosinophil activat
Human cytoskeletal
Human polynucleoti
Human cytoskeletal
Human polynucleoti
Polynucleotide enc
DNA encoding novel

DR P-PsDB; AAG91893.
 XX Novel polynucleotides derived from *Corynebacterium* bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 PS Claim 8; SEQ ID NO: 2147; 246bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the *Corynebacterium* bacterium *Corynebacterium glutamicum*. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of *Corynebacterium* bacterium, measuring expression amount and
 CC analyzing the expression profile or expression pattern of a gene derived
 CC from *Corynebacterium* bacterium, and identifying a homolog of a gene derived
 CC from *Corynebacterium* bacterium. *Corynebacterium* bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 CC
 XX
 SQ Sequence 576 BP, 101 A, 158 C, 148 G, 169 T, 0 other;
 Query Match 62.7%; Score 18.8; DB 22; Length 576;
 Best Local Similarity 90.9%; Pred. No. 1.3e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 9 AGTGTACGACCTCTGCAGGG 30
 DB 258 AGTGTACGACCTCTGCAGG 237
 AAT66463 standard; DNA, 1368 BP.
 AC AAT66463;
 XX
 DT 25-MAR-2003 (updated)
 DT 29-JUL-1997 (first entry)
 XX
 DE Thermophilic alkaline phosphatase gene.
 DE
 XX Alkaline phosphatase; thermophilic enzyme; label; assay; ss.
 XX
 OS Rhodothermus marinus strain ATCC 43812.
 OS
 XX Rhodothermus marinus strain ATCC 43812.
 XX
 FH Key Location/Qualifiers
 FH sig_peptide 1..63
 FT /tag= a
 FT /note= "the signal peptide is encoded by bases 1-60
 FT mat_peptide 64..1365
 FT /tag= b
 FT /note= "the mature protein is encoded by bases
 FT 61-1365 when the AP is expressed in *E. coli*"
 FT
 PT
 PN EP770678-A2.
 PN
 PD 02-MAY-1997.
 PD
 XX
 PF 24-OCT-1996; 96EP-0307692.
 PF
 XX
 PR 27-OCT-1995; 95US-0005965.
 PR
 XX (AMSH) AMERSHAM LIFE SCI INC.
 PA (AMSH) AMERSHAM PHARMACIA BIOTECH INC.
 XX
 PI Davis M, Szaez J;
 YX

DR WPI; 1997-238139/22.
 DR P-PsDB; AAM17830.
 XX
 PT Thermostable Rhodothermus marinus alkaline phosphatase - useful as
 PT enzyme label in immunoassays and nucleic acid assays
 PT
 XX
 PS Example; Fig 9; 24pp; English.
 XX
 CC A genomic DNA sequence (AAT66461) from *Rhodothermus marinus* (Rma)
 CC codes for a thermophilic alkaline phosphatase (AP) (AAM17830)
 CC that shows optimal activity at pH 10.8 and which is stable to
 CC heating at 65 deg for 1 hr. The sequence was deduced from a
 CC clone isolated following expression screening of *E. coli* cells
 CC transformed with Rma genomic DNA (see also AAT66461). The Rma
 CC AP gene can be used to produce recombinant AP in transformed host
 CC cells for use in nucleic acid and protein detection assays.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 XX
 SQ Sequence 1368 BP, 218 A, 451 C, 486 G, 213 T, 0 other;
 Query Match 62.7%; Score 18.8; DB 18; Length 1368;
 Best Local Similarity 76.7%; Pred. No. 1.4e+02;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 CCTGCGTAGTGTAAGACCTCTGCAGGG 30
 DB 197 CGTGTGTAAGTCTTGGCTCTGCAGCGG 226
 AAT66461 standard; DNA, 2936 BP.
 AC AAT66461;
 XX
 DT 25-MAR-2003 (updated)
 DT 29-JUL-1997 (first entry)
 XX
 DE Thermophilic alkaline phosphatase gene.
 DE
 XX Alkaline phosphatase; thermophilic enzyme; label; assay; ss.
 XX
 OS Rhodothermus marinus strain ATCC 43812.
 OS
 XX Rhodothermus marinus strain ATCC 43812.
 XX
 PN EP770678-A2.
 PN
 PD 02-MAY-1997.
 PD
 XX
 PF 24-OCT-1996; 96EP-0307692.
 PF
 XX
 PR 27-OCT-1995; 95US-0005965.
 PR
 XX (AMSH) AMERSHAM LIFE SCI INC.
 PA (AMSH) AMERSHAM PHARMACIA BIOTECH INC.
 XX
 PI Davis M, Szaez J;
 WPI; 1997-238139/22.
 PT
 PT Thermostable Rhodothermus marinus alkaline phosphatase - useful as
 PT enzyme label in immunoassays and nucleic acid assays
 PT
 XX
 PS Example; Fig 7; 24pp; English.
 XX
 CC A genomic DNA sequence (AAT66461) from *Rhodothermus marinus* (Rma)
 CC contains the gene for a thermophilic alkaline phosphatase (AP) and
 CC flanking regions. It was obtained by transforming *E. coli* XL2-Blue
 CC MRP cells with Rma genomic DNA fragments, screening for AP
 CC activity on plates using a chromogenic indicator. And KpnI
 CC digestion of isolated clone pRm1. A 1.8 kb fragment of pRm1/KpnI
 CC was subcloned into pUC19 to yield plasmid pRm1.8 (AAT66462), and
 CC a presumptive full-length coding sequence (AAT66463) for Rma AP
 CC (AAM17830) was produced. The Rma AP nucleic acids can be used for

CC prodn. of recombinant AP for use in nucleic acid and protein
 CC detection assays.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 2936 BP; 525 A; 1012 C; 880 G; 519 T; 0 other;

Query Match 62.7%; Score 18.8; DB 18; Length 2936;
 Best Local Similarity 76.7%; Pred. No. 1.5e+02;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCCTGCGTACTGCTGACGACCTCTGCGAGG 30
 DB 1351 CCCTGCGTACTGCTGCGCTCCAGCCGAG 1322

RESULT 4
 AAT66462/c
 ID AAT66462 standard; DNA; 4454 BP.

XX AAT66462;

XX 25-MAR-2003 (updated)
 DT 25-JUN-1997 (first entry)

XX Plasmid pCRM1.8 containing an alkaline phosphatase gene.

XX Alkaline phosphatase; thermophilic enzyme; label; assay;
 KW plasmid pCRM1.8; se.

XX Chimeric Rhodothermus marinus strain ATCC 43812.

XX EP70678-A2.

XX 02-MAY-1997.

XX 24-OCT-1996; 96EP-0307692.

XX 27-OCT-1995; 95US-0005965.

XX (AMSH) AMERSHAM LIFE SCI INC.
 PA (AMSH) AMERSHAM PHARMACIA BIOTECH INC.

XX Davis M, Szasz J;

XX WPI; 1997-238139/22.

XX Thermostable Rhodothermus marinus alkaline phosphatase - useful as
 PT enzyme label in immunoassays and nucleic acid assays

XX Example; Fig 8; 24pp; English.

XX Plasmid pCRM1.8 (AAT66463) contains a 1.8 kb HindIII/KpnI segment of
 CC genomic DNA (see also AAT66461) containing the alkaline phosphatase
 CC (AP) and flanking sequences from Rhodothermus marinus (Rma), cloned
 CC into pUC19. This recombinant plasmid has been used to produce Rma
 CC thermophilic AP in E. coli JM109 transformants. A presumptive
 CC full-length coding sequence (AAT66463) for Rma AP (AAW17830) was
 CC deduced. The Rma AP nucleic acids can be used for prodn. of
 CC recombinant AP for use in nucleic acid and protein detection
 CC assays.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 4454 BP; 978 A; 1305 C; 1234 G; 937 T; 0 other;

Query Match 62.7%; Score 18.8; DB 18; Length 4454;
 Best Local Similarity 76.7%; Pred. No. 1.5e+02;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCCTGCGTACTGCTGACGACCTCTGCGAGG 30
 DB 1738 CCCTGCGTACTGCTGCGCTCCAGCCGAG 1709

RESULT 5
 AAH68530/c
 ID AAH68530 standard; DNA; 34980 BP.

XX AAH68530;

XX 26-SEP-2001 (first entry)

XX C glutamicum coding sequence fragment SEQ ID NO: 7065.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; de.

XX Corynebacterium glutamicum.

XX EP108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-0127688.

XX 16-DEC-1999; 99JP-0377484.

XX 07-APR-2000; 2000JP-0159162.

XX 03-AUG-2000; 2000JP-0280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H,
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PS
 XX Disclosure; SEQ ID NO: 7065; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.

XX Sequence 349980 BP; 86896 A; 98023 C; 80939 G; 84122 T; 0 other;

Query Match 62.7%; Score 18.8; DB 22; Length 349980;
 Best Local Similarity 90.9%; Pred. No. 2.1e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 AGGTGACGACCTCTGCGAGG 30
 DB 266823 AGGTGACGACCTCTGCGAGG 266802

RESULT 6
 ABN17942
 ID ABN17942 standard; cDNA; 294 BP.

XX ABN17942;

XX 24-JUN-2002 (first entry)

XX Human ORF polynucleotide sequence SEQ ID NO:4361.

XX	Human, open reading frame; ORF; gene therapy; cancer; cirrhosis;
KM	hyperproliferative disorder; porriasis; benign tumour; haemorrhage;
KM	degenerative disorder; osteoarthritis; neurodegenerative disorder;
KM	cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KM	hypertension; hypothyroidism; cholesterol ester storage disease;
KM	immune deficiency; immune disorder; infectious disease;
KM	autoimmune disease; rheumatoid arthritis; autoimmune thyroiditis;
KM	myasthenia gravis; gene; 88.
XX	
OS	Homo sapiens.
XX	
PN	WO200192523-A2.
XX	
PD	06-DEC-2001.
XX	
FP	29-MAY-2001, 2001WO-US10636.
XX	
XX	30-MAY-2000, 2000US-206132P.
PR	29-AUG-2000, 2000US-228716P.
XX	
PA	(CURA-) CURAGEN CORP.
XX	
PI	Shinkets RA, Leach MD;
XX	
DR	WPI, 2002-106308/14.
XX	
DR	P-PSDB; ABP02190.
XX	
PT	Novel human polypeptides and polynucleotides useful for diagnosing,
PT	preventing and treating cardiovascular disease, neurodegenerative,
PT	hyperproliferative disorders and autoimmune disorders -
XX	
XX	Disclosure, SEQ ID 4361, 1037pp; English.

CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification), AB151762 to AB272752 encode the human ORFX
 CC proteins given in ABP00010 to ABP1500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal diseases, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
 CC
 CC Sequence 294 BP; 91 A; 63 C; 82 G; 56 T; 2 other;

[illegible]

RESULT 7
JBI99918

```

ID  AB1999218 standard; cDNA; 3653 BP.
XX
AC  AB1999218;
XX
DT  08-MAR-2002 (first entry)
XX
DE  Rat mucocardial cell proliferation associated cDNA SEQ ID NO 7.
XX
KW  Rat; heart; cardiac; myocardial necrosis; cardiac hypertrophy;
KM  cardiac insufficiency; ss.
XX
OS  Rattus norvegicus.
XX
PN  W0200183705-A1.
XX
PD  08-NOV-2001.
XX
PF  27-APR-2001; 2001WO-JP03700.
XX
PR  27-APR-2000; 2000JP-0126741.
XX
PA  (KYOM ) KYOMA HAKKO KOGYO KK.
XX
PI  Yamada Y, Sekine S, Kikuchi Y, Sakurada K;
XX
XX  MPI; 2002-075160/10.
DR  P-PSDB; ABB57378.
XX
PT  Genes having differential expression in fetal and adult heart tissue
PT  useful for screening potential drugs for promoting repair of damage
PT  caused by myocardial necrosis
XX
PS  Claim 22; Page 84-90; 171pp; Japanese.

```

The invention relates to gene sequences (ABJ199915-ABJ199934) having modified expression in fetal heart tissue as compared to adult heart tissue and the encoded proteins (ABB57375-ABB57392). The genes have cardiac activity and may be useful in the promotion of the repair of damage to heart tissue caused by myocardial necrosis. The gene sequences are useful for screening potential compounds for the ability to influence disease associated with myocardial necrosis. Drugs identified by the screening methods may be used to treat and prevent disease with which myocardial necrosis is associated, such as cardiac hypertrophy and cardiac insufficiency. Diagnosis of diseases such as those above is also disclosed.

CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC

Sequence 3653 BP; 845 A; 1090 C; 1048 G; 670 T; 0 other:
SQ

Query Match 61.3%; Score 18.4; DB 24; Length 3653;
Best Local Similarity 78.6%; Pred. No. 2,3e+02;
Matches 22; conservative 0; Mismatches 6; Indels 0; Gaps 0

CY 2 CCTGGCTAGTGTATGACACTCCTGCAGC 29
DB 1788 CCTGGCTATTCTTTTGCGGCTCTCTGCATG 1815

	Query Match	Similarity	61.3%	Score 18.4	DB 24	Length 3553
	Best Local	Matches	78.6%	Pident No. 2.3e+02	Mismatches 6	Indels 0; Gaps 0
Oy	2	CCTGCGTAGTGTACGACCCTCCTGAGC	29			
Db	1788	CCTGCGTTCTTGTCGGCTCCTGCATG	1815			
RESULT 8						
ID	ABA96146					
XX	ABA96146	standard; DNA; 4705 BP.				
AC	ABA96146;					
XX	15-APR-2002	(first entry)				
DT						
XX						
DE	Modified human/mouse Bcr-Abl fusion gene #1.					
XX						
KX	Human; mouse; Bcr-Abl; cancer; cytostatic; immunosuppressive; leukaemia;					
KM	orthologous bone marrow transplantation; ds.					

OS	Chimeric - <i>Mus musculus</i> .
XX	

```

FH Key Location/Qualifiers
FT CDS 1..4704
FT /tag= a
FT /product= "Bcr-Abl fusion protein #1"
XX
XX WO200200024-A1.
XX
XX 03-JAN-2002.
XX
XX 29-JUN-2001; 2001WO-US20602.
XX
XX 30-JUN-2000; 2000US-215595P.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Wang JYJ, Vigneri P;
XX
XX WPI; 2002-139846/18.
XX
XX P-PSDB; ABB08236.
XX
XX Killing cancer cells for treating leukaemia, involves causing
XX accumulation of activated Bcr-Abl in the nucleus to induce apoptosis -
XX
XX Disclosure; Page 33-40; 63pp; English.
XX
XX The sequence encodes a human/mouse Bcr-Abl fusion protein. The invention
XX relates to a novel method for killing cancer cells, comprising
XX causing an accumulation of activated Bcr-Abl in the nucleus, to induce
XX apoptosis. The method has cytostatic and immunosuppressive activity. The
XX method is useful for killing cancer cells, and in the treatment of
XX cancers including leukaemia. The invention also supplies a method useful
XX for purging bone marrow to allow for orthologous bone marrow
XX transplantation, and as a therapeutic option for chronic myelogenous
XX leukaemia.
XX
SQ Sequence 4705 BP; 1072 A; 1457 C; 1435 G; 741 T; 0 other;
XX
XX Query Match 61.3%; Score 18.4; DB 24; Length 4705;
XX Best Local Similarity 78.6%; Pred. No. 2.3e+02;
XX Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
OY 2 CCTGCGTAGTGTGACGACCTCTGCAGG 29
DB 3414 CCTGCGTCTGTGTCGGCCTCTGCATG 3441
XX
XX RESULT 9
XX ABA96147 standard; DNA; 4786 BP.
XX
XX ABA96147;
XX
XX 15-APR-2002 (first entry)
XX
XX Modified human/mouse Bcr-Abl fusion gene #2.
XX
XX Human; mouse; Bcr-Abl; cancer; cytostatic; immunosuppressive; leukaemia;
XX orthologous bone marrow transplantation; ds.
XX
XX Chimeric - Homo sapiens.
XX
XX OS Chimeric - Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 1..4785
XX /tag= a
XX /product= "Bcr-Abl fusion protein #2"
XX
XX WO200200024-A1.
XX
XX 03-JAN-2002.
XX
XX 29-JUN-2001; 2001WO-US20602.
XX

```

```

PR 30-JUN-2000; 2000US-215595P.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Wang JYJ, Vigneri P;
XX
XX WPI; 2002-139846/18.
XX
XX P-PSDB; ABB08237.
XX
XX Killing cancer cells for treating leukaemia, involves causing
XX accumulation of activated Bcr-Abl in the nucleus to induce apoptosis -
XX
XX Example 10; Page 47-54; 63pp; English.
XX
XX The sequence encodes a human/mouse Bcr-Abl fusion protein. The invention
XX relates to a novel method for killing cancer cells, comprising
XX causing an accumulation of activated Bcr-Abl in the nucleus, to induce
XX apoptosis. The method has cytostatic and immunosuppressive activity. The
XX method is useful for killing cancer cells, and in the treatment of
XX cancers including leukaemia. The invention also supplies a method useful
XX for purging bone marrow to allow for orthologous bone marrow
XX transplantation, and as a therapeutic option for chronic myelogenous
XX leukaemia.
XX
SQ Sequence 4786 BP; 1104 A; 1470 C; 1465 G; 747 T; 0 other;
XX
XX Query Match 61.3%; Score 18.4; DB 24; Length 4786;
XX Best Local Similarity 78.6%; Pred. No. 2.3e+02;
XX Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
OY 2 CCTGCGTAGTGTGACGACCTCTGCAGG 29
DB 3414 CCTGCGTCTGTGTCGGCCTCTGCATG 3441
XX
XX RESULT 10
XX ABT00010/c
XX ID ABT00010 standard; DNA; 1503841 BP.
XX
XX ABT00010;
XX
XX 07-NOV-2002 (first entry)
XX
XX Human neurogulin 1 gene.
XX
XX Human; neurogulin 1; neurogulin-1-associated gene 1; NRGL1;
XX schizophrenia; chromosome 8p12; single nucleotide polymorphism; SNP;
XX neuroleptic; gene therapy; gene; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 244312..1369465
XX /tag= a
XX /product= "neurogulin 1"
XX /note= "this sequence contains introns"
XX /tag= b
XX /number= 1
XX /tag= c
XX /number= 2
XX /tag= d
XX /number= 2
XX /tag= e
XX /number= 2
XX /tag= f
XX /number= 3
XX /tag= g
XX
XX exon
XX
XX intron
XX
XX intron
XX
XX intron
XX

```

FT /number= 3
 FT exon 826308..826355
 FT /tag= h
 FT /number= 4
 FT intron 826356..1034242
 FT /tag= i
 FT /number= 4
 FT exon 1034243..1034321
 FT /tag= j
 FT /number= 5
 FT intron 1034322..1153294
 FT /tag= k
 FT /number= 5
 FT exon 1153295..1153886
 FT /tag= l
 FT /number= 6
 FT intron 1153887..1200721
 FT /tag= m
 FT /number= 6
 FT exon 1200722..1201065
 FT /tag= n
 FT /number= 7a
 FT intron 1153887..1200887
 FT /tag= o
 FT /number= 7a
 FT exon 1200888..1201065
 FT /tag= p
 FT /number= 7b
 FT intron 1201066..1210622
 FT /tag= q
 FT /number= 7b
 FT exon 1210623..1210744
 FT /tag= r
 FT /number= 8
 FT intron 1210745..1219542
 FT /tag= s
 FT /number= 8
 FT exon 1219543..1219593
 FT /tag= t
 FT /number= 9
 FT intron 1219594..1221863
 FT /tag= u
 FT /number= 9
 FT exon 1221864..1221914
 FT /tag= v
 FT /number= 10
 FT intron 1221915..1252253
 FT /tag= w
 FT /number= 10
 FT exon 1252254..1253413
 FT /tag= x
 FT /number= 11
 FT intron 1253414..1326781
 FT /tag= y
 FT /number= 11
 FT exon 1326782..1327071
 FT /tag= z
 FT /number= 12
 FT intron 1327072..1332977
 FT /tag= aa
 FT /number= 12
 FT exon 1332978..1333107
 FT /tag= ab
 FT /number= 13a
 FT intron 1333108..1347039
 FT /tag= ac
 FT /number= 13a
 FT exon 1332978..1333652
 FT /tag= ad
 FT /number= 13b
 FT intron 1333653..1347039
 FT /tag= ae
 FT /number= 13b
 FT

FT 1347040..1347107
 FT exon /tag= af
 FT /number= 14
 FT intron 1347108..1347706
 FT /tag= ag
 FT /number= 14
 FT exon 1347707..1347765
 FT /tag= ah
 FT /number= 15a
 FT intron 1347766..1354620
 FT /tag= aj
 FT /number= 15a
 FT exon 1347707..1348257
 FT /tag= ak
 FT /number= 15b
 FT intron 1348258..1354620
 FT /tag= al
 FT /number= 15b
 FT exon 1354621..1354644
 FT /tag= am
 FT /number= 16
 FT intron 1354645..1359431
 FT /tag= an
 FT /number= 16
 FT exon 1359432..1359534
 FT /tag= ao
 FT /number= 17
 FT intron 1359535..1361462
 FT /tag= ap
 FT /number= 17
 FT exon 1361463..1361589
 FT /tag= aq
 FT /number= 18
 FT intron 1361590..1364362
 FT /tag= ar
 FT /number= 18
 FT exon 1364363..1364493
 FT /tag= as
 FT /number= 19
 FT intron 1364494..1365256
 FT /tag= at
 FT /number= 19
 FT exon 1365257..1365463
 FT /tag= au
 FT /number= 20a
 FT intron 1365464..1368280
 FT /tag= av
 FT /number= 20a
 FT exon 1365257..1366044
 FT /tag= aw
 FT /number= 20b
 FT intron 1366045..1368280
 FT /tag= ax
 FT /number= 20b
 FT exon 1368281..1368422
 FT /tag= ay
 FT /number= 21
 FT intron 1368423..1368810
 FT /tag= az
 FT /number= 21
 FT exon 1368811..1369656
 FT /tag= ba
 FT /number= 22
 FT CDS 629728..744437
 FT /tag= bb
 FT /product= "neuregulin-1-associated gene 1"
 FT /note= "this sequence contains introns"
 FT 629728..629804
 FT /tag= bc
 FT /number= 1
 FT intron 629805..630974
 FT /tag= bd
 FT /number= 1
 FT

FT exon 630975..631187
FT /tag= be
FT /number= 2
FT intron 631188..631282
FT /tag= bf
FT /number= 2
FT exon 631283..631320
FT /tag= bg
FT /number= 3
FT intron 631321..634334
FT /tag= bh
FT exon 634335..634441
FT /tag= bi
FT /number= 4a
FT intron 634442..635331
FT /tag= bj
FT /number= 4a

Query Match 61.3%; Score 18.4; DB 24; Length 1503841;
Best Local Similarity 78.6%; Pred. No. 3.1e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CCTGCCGACGTGCTACGACCTCCTGCAG 28
Db 801369 CCTGACTAGTGGAGACAGCCTGCAG 801342

RESULT 11

ABT01503/c
ID ABT01503 standard; DNA; 1503841 BP.

XX AC ABT01503;

DT 07-NOV-2002 (first entry)

XX DE Human neuregulin 1 gene.

XX KW Human; neuregulin 1; neuregulin-1-associated gene 1; NRGL; NRGL1;

KW schizophrenia; chromosome 8p12; single nucleotide polymorphism; SNP;

XX KW neuroleptic; gene therapy; gene; ds.

XX OS Homo sapiens.

XX Key

FT CDS

FT .

FT exon

FT intron

FT intron

FT exon

FT intron

FT intron

FT exon

FT intron

FT exon

FT intron

FT intron

FT exon 1034243..1034321
FT /tag= j
FT /number= 5
FT intron 1034322..1153294
FT /tag= k
FT /number= 5
FT exon 1153295..1153886
FT /tag= l
FT /number= 6
FT intron 1153887..1200721
FT /tag= m
FT /number= 6
FT exon 1200722..1201065
FT /tag= n
FT /number= 7a
FT intron 1153887..1200887
FT /tag= o
FT /number= 7a
FT exon 1200888..1201065
FT /tag= p
FT /number= 7b
FT intron 1201066..1210622
FT /tag= q
FT /number= 7b
FT exon 1210623..1210744
FT /tag= r
FT /number= 8
FT intron 1210745..1219542
FT /tag= s
FT /number= 8
FT exon 1219543..1219593
FT /tag= t
FT /number= 9
FT intron 1219594..1221863
FT /tag= u
FT /number= 9
FT exon 1221864..1221914
FT /tag= v
FT /number= 10
FT intron 1221915..1252253
FT /tag= w
FT /number= 10
FT exon 1252254..1253413
FT /tag= x
FT /number= 11
FT intron 1253414..1326781
FT /tag= y
FT /number= 11
FT exon 1326782..1327071
FT /tag= z
FT /number= 12
FT intron 1327072..1332977
FT /tag= aa
FT /number= 12
FT exon 1332978..1333107
FT /tag= ab
FT /number= 13a
FT intron 1333108..1347039
FT /tag= ac
FT /number= 13a
FT exon 1332978..1333652
FT /tag= ad
FT /number= 13b
FT intron 1333653..1347039
FT /tag= ae
FT /number= 13b
FT exon 1347040..1347107
FT /tag= af
FT /number= 14
FT intron 1347108..1347706
FT /tag= ag
FT /number= 14
FT exon 1347707..1347765

FT	/*cag= ah	/number= 15a
FT	1347766..1354620	
FT	/*cag= aj	/number= 15a
FT	1347707..1348257	
FT	/*cag= ak	/number= 15b
FT	1348258..1354620	
FT	/*cag= al	/number= 15b
FT	1354621..1354644	
FT	/*cag= am	/number= 16
FT	1354645..1359431	
FT	/*cag= an	/number= 16
FT	1359432..1359534	
FT	/*cag= ao	/number= 17
FT	1361590..1364362	
FT	/*cag= ar	/number= 18
FT	1364363..1364493	
FT	/*cag= as	/number= 19
FT	1364494..1365256	
FT	/*cag= at	/number= 19
FT	1365257..1365463	
FT	/*cag= au	/number= 20a
FT	1365464..1368280	
FT	/*cag= av	/number= 20a
FT	1365257..1366044	
FT	/*cag= aw	/number= 20b
FT	1366045..1368280	
FT	/*cag= ax	/number= 20b
FT	1368281..1368422	
FT	/*cag= ay	/number= 21
FT	1368423..1368810	
FT	/*cag= az	/number= 21
FT	1368811..1369656	
FT	/*cag= ba	/number= 22
FT	629728..74437	
FT	/*cag= bb	/product= "neuregulin-1-associated gene 1"
FT	/*cag= bc	/note= "this sequence contains introns"
FT	/*cag= bd	/number= 1
FT	629805..630974	
FT	/*cag= bd	/number= 1
FT	630975..631187	
FT	/*cag= be	/number= 2
FT	631188..631282	
FT	/*cag= bf	/number= 2
FT	631283..631320	
FT	/*cag= bh	/number= 15a
FT	1347766..1354620	
FT	/*cag= aj	/number= 15a
FT	1347707..1348257	
FT	/*cag= ak	/number= 15b
FT	1348258..1354620	
FT	/*cag= al	/number= 15b
FT	1354621..1354644	
FT	/*cag= am	/number= 16
FT	1354645..1359431	
FT	/*cag= an	/number= 16
FT	1359432..1359534	
FT	/*cag= ao	/number= 17
FT	1361590..1364362	
FT	/*cag= ar	/number= 18
FT	1364363..1364493	
FT	/*cag= as	/number= 19
FT	1364494..1365256	
FT	/*cag= at	/number= 19
FT	1365257..1365463	
FT	/*cag= au	/number= 20a
FT	1365464..1368280	
FT	/*cag= av	/number= 20a
FT	1365257..1366044	
FT	/*cag= aw	/number= 20b
FT	1366045..1368280	
FT	/*cag= ax	/number= 20b
FT	1368281..1368422	
FT	/*cag= ay	/number= 21
FT	1368423..1368810	
FT	/*cag= az	/number= 21
FT	1368811..1369656	
FT	/*cag= ba	/number= 22
FT	629728..74437	
FT	/*cag= bb	/product= "neuregulin-1-associated gene 1"
FT	/*cag= bc	/note= "this sequence contains introns"
FT	/*cag= bd	/number= 1
FT	629805..630974	
FT	/*cag= bd	/number= 1
FT	630975..631187	
FT	/*cag= be	/number= 2
FT	631188..631282	
FT	/*cag= bf	/number= 2
FT	631283..631320	
FT	/*cag= bh	/number= 15a
FT	1347766..1354620	
FT	/*cag= aj	/number= 15a
FT	1347707..1348257	
FT	/*cag= ak	/number= 15b
FT	1348258..1354620	
FT	/*cag= al	/number= 15b
FT	1354621..1354644	
FT	/*cag= am	/number= 16
FT	1354645..1359431	
FT	/*cag= an	/number= 16
FT	1359432..1359534	
FT	/*cag= ao	/number= 17
FT	1361590..1364362	
FT	/*cag= ar	/number= 18
FT	1364363..1364493	
FT	/*cag= as	/number= 19
FT	1364494..1365256	
FT	/*cag= at	/number= 19
FT	1365257..1365463	
FT	/*cag= au	/number= 20a
FT	1365464..1368280	
FT	/*cag= av	/number= 20a
FT	1365257..1366044	
FT	/*cag= aw	/number= 20b
FT	1366045..1368280	
FT	/*cag= ax	/number= 20b
FT	1368281..1368422	
FT	/*cag= ay	/number= 21
FT	1368423..1368810	
FT	/*cag= az	/number= 21
FT	1368811..1369656	
FT	/*cag= ba	/number= 22
FT	629728..74437	
FT	/*cag= bb	/product= "neuregulin-1-associated gene 1"
FT	/*cag= bc	/note= "this sequence contains introns"
FT	/*cag= bd	/number= 1
FT	629805..630974	
FT	/*cag= bd	/number= 1
FT	630975..631187	
FT	/*cag= be	/number= 2
FT	631188..631282	
FT	/*cag= bf	/number= 2
FT	631283..631320	
FT	/*cag= bh	/number= 15a
FT	1347766..1354620	
FT	/*cag= aj	/number= 15a
FT	1347707..1348257	
FT	/*cag= ak	/number= 15b
FT	1348258..1354620	
FT	/*cag= al	/number= 15b
FT	1354621..1354644	
FT	/*cag= am	/number= 16
FT	1354645..1359431	
FT	/*cag= an	/number= 16
FT	1359432..1359534	
FT	/*cag= ao	/number= 17
FT	1361590..1364362	
FT	/*cag= ar	/number= 18
FT	1364363..1364493	
FT	/*cag= as	/number= 19
FT	1364494..1365256	
FT	/*cag= at	/number= 19
FT	1365257..1365463	
FT	/*cag= au	/number= 20a
FT	1365464..1368280	
FT	/*cag= av	/number= 20a
FT	1365257..1366044	
FT	/*cag= aw	/number= 20b
FT	1366045..1368280	
FT	/*cag= ax	/number= 20b
FT	1368281..1368422	
FT	/*cag= ay	/number= 21
FT	1368423..1368810	
FT	/*cag= az	/number= 21
FT	1368811..1369656	
FT	/*cag= ba	/number= 22
FT	629728..74437	
FT	/*cag= bb	/product= "neuregulin-1-associated gene 1"
FT	/*cag= bc	/note= "this sequence contains introns"
FT	/*cag= bd	/number= 1
FT	629805..630974	
FT	/*cag= bd	/number= 1
FT	630975..631187	
FT	/*cag= be	/number= 2
FT	631188..631282	
FT	/*cag= bf	/number= 2
FT	631283..631320	
FT	/*cag= bh	/number= 15a
FT	1347766..1354620	
FT	/*cag= aj	/number= 15a
FT	1347707..1348257	
FT	/*cag= ak	/number= 15b
FT	1348258..1354620	
FT	/*cag= al	/number= 15b
FT	1354621..1354644	
FT	/*cag= am	/number= 16
FT	1354645..1359431	
FT	/*cag= an	/number= 16
FT	1359432..1359534	
FT	/*cag= ao	/number= 17
FT	1361590..1364362	
FT	/*cag= ar	/number= 18
FT	1364363..1364493	
FT	/*cag= as	/number= 19
FT	1364494..1365256	
FT	/*cag= at	/number= 19
FT	1365257..1365463	
FT	/*cag= au	/number= 20a
FT	1365464..1368280	
FT	/*cag= av	/number= 20a
FT	1365257..1366044	
FT	/*cag= aw	/number= 20b
FT	1366045..1368280	
FT	/*cag= ax	/number= 20b
FT	1368281..1368422	
FT	/*cag= ay	/number= 21
FT	1368423..1368810	
FT	/*cag= az	/number= 21
FT	1368811..1369656	
FT	/*cag= ba	/number= 22
FT	629728..74437	
FT	/*cag= bb	/product= "neuregulin-1-associated gene 1"
FT	/*cag= bc	/note= "this sequence contains introns"
FT	/*cag= bd	/number= 1
FT	629805..630974	
FT	/*cag= bd	/number= 1
FT	630975..631187	
FT	/*cag= be	/number= 2
FT	631188..631282	
FT	/*cag= bf	/number= 2
FT	631283..631320	
FT	/*cag= bh	/number= 15a
FT	1347766..1354620	
FT	/*cag= aj	/number= 15a
FT	1347707..1348257	
FT	/*cag= ak	/number= 15b
FT	1348258..1354620	
FT	/*cag= al	/number= 15b
FT	1354621..1354644	
FT	/*cag= am	/number= 16
FT	1354645..1359431	
FT	/*cag= an	/number= 16
FT	1359432..1359534	
FT	/*cag= ao	/number= 17
FT	1361590..1364362	
FT	/*cag= ar	/number= 18
FT	1364363..1364493	
FT	/*cag= as	/number= 19
FT	1364494..1365256	
FT	/*cag= at	/number= 19
FT	1365257..1365463	
FT	/*cag= au	/number= 20a
FT	1365464..1368280	
FT	/*cag= av	/number= 20a
FT	1365257..1366044	
FT	/*cag= aw	/number= 20b
FT	1366045..1368280	
FT	/*cag= ax	/number= 20b
FT	1368281..1368422	
FT	/*cag= ay	/number= 21
FT	1368423..1368810	
FT	/*cag= az	/number= 21
FT	1368811..1369656	
FT	/*cag= ba	/number= 22
FT	629728..74437	
FT	/*cag= bb	/product= "neuregulin-1-associated gene 1"
FT	/*cag= bc	/note= "this sequence contains introns"
FT	/*cag= bd	/number= 1
FT	629805..630974	
FT	/*cag= bd	/number= 1
FT	630975..631187	
FT	/*cag= be	/number= 2
FT	631188..631282	
FT	/*cag= bf	/number= 2
FT	631283..631320	
FT	/*cag= bh	/number= 15a
FT	1347766..1354620	
FT	/*cag= aj	/number= 15a
FT	1347707..1348257	
FT	/*cag= ak	/number= 15b
FT	1348258..1354620	
FT	/*cag= al	/number= 15b
FT	1354621..1354644	
FT	/*cag= am	/number= 16
FT	1354645..1359431	
FT	/*cag= an	/number= 16
FT	1359432..1359534	
FT	/*cag= ao	/number= 17
FT	1361590..1364362	
FT	/*cag= ar	/number= 18
FT	1364363..1364493	
FT	/*cag= as	/number= 19
FT	1364494..1365256	
FT	/*cag= at	/number= 19
FT	1365257..1365463	
FT	/*cag= au	/number= 20a
FT	1365464..1368280	
FT	/*cag= av	/number= 20a
FT	1365257..1366044	
FT	/*cag= aw	/number= 20b
FT	1366045..1368280	
FT	/*cag= ax	/number= 20b
FT	1368281..1368422	
FT	/*cag= ay	/number= 21
FT	1368423..1368810	
FT	/*cag= az	/number= 21
FT	1368811..1369656	
FT	/*cag= ba	/number= 22
FT	629728..74437	
FT	/*cag= bb	/product= "neuregulin-1-associated gene 1"
FT	/*cag= bc	/note= "this sequence contains introns"
FT	/*cag= bd	/number= 1
FT	629805..630974	
FT	/*cag= bd	/number= 1
FT	630975..631187	
FT	/*cag= be	/number= 2
FT	631188..631282	
FT	/*cag= bf	/number= 2
FT	631283..631320	
FT	/*cag= bh	/number= 15a
FT	1347766..1354620	
FT	/*cag= aj	/number= 15a
FT	1347707..1348257	
FT	/*cag= ak	/number= 15b
FT	1348258..1354620	
FT	/*cag= al	/number= 15b
FT	1354621..1354644	
FT	/*cag= am	/number= 16
FT	1354645..1359431	
FT	/*cag= an	/number= 16
FT	1359432..1359534	
FT	/*cag= ao	/number= 17
FT	1361590..1364362	
FT	/*cag= ar	/number= 18
FT	1364363..1364493	
FT	/*cag= as	/number= 19
FT	1364494..1365256	
FT	/*cag= at	/number= 19
FT	1365257..1365463	
FT	/*cag= au	/number= 20a
FT	1365464..1368280	
FT	/*cag= av	/number= 20a
FT	1365257..1366044	
FT	/*cag= aw	/number= 20b
FT	1366045..1368280	
FT	/*cag= ax	/number= 20b
FT	1368281..1368422	
FT	/*cag= ay	/number= 21
FT	1368423..1368810	
FT	/*cag= az	/number= 21
FT	1368811..1369656	
FT	/*cag= ba	/number= 22
FT	629728..74437	
FT	/*cag= bb	/product= "neuregulin-1-associated gene 1"
FT	/*cag= bc	/note= "this sequence contains introns"
FT	/*cag= bd	/number= 1
FT	629805..630974	
FT	/*cag= bd	/number= 1
FT	630975..631187	
FT	/*cag= be	/number= 2
FT	631188..631282	
FT	/*cag= bf	/number= 2
FT	631283..631320	
FT	/*cag= bh	/number= 15a
FT	1347766..1354620	
FT	/*cag= aj	/number= 15a
FT	1347707..1348257	
FT	/*cag= ak	/number= 15b
FT	1348258..1354620	
FT	/*cag= al	/number= 15b
FT	1354621..1354644	
FT	/*cag= am	/number= 16
FT	1354645..1359431	
FT	/*cag= an	/number= 16
FT	1359432..1359534	
FT	/*cag= ao	/number= 17
FT	1361590..1364362	
FT	/*cag= ar	/number= 18
FT	1364363..1364493	
FT	/*cag= as	/number= 19
FT	1364494..1365256	
FT	/*cag= at	/number= 19
FT	1365257..1365463	
FT	/*cag= au	/number= 20a
FT	1365464..1368280	
FT	/*cag= av	/number= 20a
FT	1365257..1366044	
FT	/*cag= aw	/number= 20b
FT	1366045..1368280	
FT	/*cag= ax	/number= 20b
FT	1368281..1368422	
FT	/*cag= ay	/number= 21
FT	1368423..1368810	
FT	/*cag= az	/number= 21
FT	1368811..1369656	
FT	/*cag= ba	/number= 22
FT	629728..74437	
FT	/*cag= bb	/product= "neure

```

FT      /*tag= bg
FT      /number= 3
FT      631321..634334
FT      /tag= bh
FT      /number= 3
FT      634335..634441
FT      /*tag= b1
FT      /number= 4a
FT      634442..635331
FT      /tag= bj
FT      /number= 4a

Query Match      61.3%; Score 18.4; DB 24; Length 1503841;
Basic Local Similarity 78.6%; Pred. No. 3,1e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0

Db      1 CCCTGCGTAGTGATACGACCTCCCTGCAG 28
      601369 CCCTGACTAGTGGGAGACGACCTCTGCAG 801342

RESULT 12
ID      AAK95240/c
XX      AAK95240 standard; DNA; 1503900 BP.
XX
XX      AAK95240;
XX
XX      17-DEC-2001 (first entry)
XX
XX      Human neuregulin-1 gene.
XX      DE
XX      Human neuregulin-1 gene.
XX      KW
XX      Human; neuregulin-1 associated gene 1; NRGIAG1; Schizophrenia gene;
XX      KM
XX      gene therapy; ds.
XX
XX      OS
XX      Homo sapiens.
XX
XX      WC200164876-A2.
XX
XX      PD
XX      07-SEP-2001.
XX
XX      28-FEB-2001; 2001WO-US06376.
XX      PF
XX      28-FEB-2000; 2000US-0515715.
XX      XX
XX      (DECO-) DECODE GENETICS EHF;
XX      PA
XX      Stefanesson H, Steinhorsdottir V, Gulcher JR;
XX      PI
XX      WPI; 2001-550179/61.
XX      DR
XX      P-PSDI; AAG67900, AAG67901, AAG67902, AAG67903, AAG67904, AAG67905,
XX      DR      AAG67906, AAG67907, AAG67908, AAG67909, AAG67910, AAG67911, AAG67912,
XX      DR      AAG67913, AAG67914, AAG67915, AAG67916, AAG67918, AAG67919,
XX      DR      AAG67920, AAG67921, AAG67922, AAG67923, AAG67924, AAG67926,
XX      DR      AAG67927, AAG67928, AAG67929, AAG67930, AAG67931, AAG67932, AAG67933,
XX      DR      AAG67934, AAG67935, AAG67936, AAG67937.
XX      XX
XX      Neuregulin-1 associated gene 1 nucleic acids and fragments, useful for
XX      PT      preventing diagnosing and treating schizophrenia -
XX      PS      Disclosure; Page 90-501; 750bp; English.
XX
XX      This sequence represents the human neuregulin-1 associated gene 1
XX      CC      (NRGIAG1) of the invention. The NRGIAG1 gene is also referred to as the
XX      CC      human Schizophrenia gene. The invention also relates to fragments or
XX      CC      variants of the gene and the NRGIAG1 polypeptides they encode. The
XX      CC      NRGIAG1 nucleic acids and polypeptides may be used in the prevention,
XX      CC      diagnosis and treatment of diseases associated with inappropriate NRGIAG1
XX      CC      expression. For example, they may be used to treat disorders associated
XX      CC      with decreased expression by rectifying mutations or deletions in a
XX      CC      patient's genome that affect the activity of NRGIAG1 by expressing
XX      CC      inactive proteins or to supplement the patients own production of
XX      CC      NRGIAG1. Additionally, the gene may be used to produce NRGIAG1
XX      CC      polypeptides by inserting the nucleic acids into a host cell and

```

CC culturing the cell to express the protein. The gene may also be used as
 CC DNA probes and primers in diagnostic assays to detect and quantitate the
 CC presence of similar nucleic acids in samples, and therefore which
 CC patients may be in need of restorative therapy. The NRG1AG1 polypeptides
 CC may also be used as antigens in the production of antibodies against
 CC NRG1AG1 and in assays to identify modulators of NRG1AG1 expression and
 CC activity. Anti-NRG1AG1 antibodies and antagonists may also be used to
 CC down regulate expression and activity. Anti-NRG1AG1 antibodies may
 CC also be used as diagnostic agents for detecting the presence of NRG1AG1
 CC polypeptides in samples. NRG1AG1 is associated with schizophrenia which
 CC may be prevented, diagnosed and/or treated by the above methods.

SO Sequence 1503900 BP; 452487 A; 281874 C; 288074 G; 480092 T; 1373 other;

Query Match 61.3%; Score 18.4; DB 22; Length 1503900;
 Best Local Similarity 78.6%; Pred. No. 3.1e+02;
 Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CCTGTGCTAGTGTGACGACCTCTGCAG 28
 Db 801368 CCTGTGCTAGTGTGACGACGACCTCTGCAG 801341

RESULT 13
 AAK96733/C
 ID AAK96733 standard; DNA; 1503900 BP.

XX AAK96733;

DT 17-DEC-2001 (first entry)

DE Human neuregulin-1 gene.

XX Human; neuregulin 1 gene; schizophrenia; gene therapy; de.

OS Homo sapiens.

XX MO200164877-A2.

PD 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US06377.

PR 28-FEB-2000; 2000US-0515716.

XX (DECO-) DECODE GENETICS EHF.

PA Stefansson H, Steinthorsdottir V, Gulcher JR;

XX WPI; 2001-514841/56.

DR P-PSDB; AAG67938, AAG67939, AAG67940, AAG67941, AAG67942, AAG67943,

DR AAG67944, AAG67945, AAG67946, AAG67947, AAG67948, AAG67949,

DR AAG67950, AAG67951, AAG67952, AAG67953, AAG67954, AAG67955,

DR AAG67956, AAG67957, AAG67958, AAG67959, AAG67960, AAG67961,

DR AAG67962, AAG67963, AAG67964, AAG67965, AAG67966, AAG67967,

DR AAG67968, AAG67969, AAG67970, AAG67971, AAG67972, AAG67973,

DR AAG67974, AAG67975.

XX Neuregulin 1 nucleic acids and proteins useful for diagnosing

PT preventing and treating schizophrenia -

XX Disclosure; Page 345-756; 756pp; English.

CC This sequence represents the human neuregulin 1 gene of the invention.
 CC The invention also relates to fragments or variants of the neuregulin 1
 CC gene. The gene and its proteins may be used in the prevention, diagnosis
 CC and treatment of diseases associated with inappropriate neuregulin 1
 CC expression, such as schizophrenia. For example they may be used to treat
 CC disorders associated with decreased neuregulin 1 expression by rectifying
 CC mutations or deletions in a patient's genome that affect the activity of
 CC neuregulin 1 by expressing inactive proteins or to supplement the
 CC patients own production of polypeptides. Additionally, the gene may be
 CC used to produce the neuregulin 1 protein, by inserting the nucleic acids

CC into a host cell and culturing the cell to express the protein. The gene
 CC and its complementary sequences may also be used as DNA probes in
 CC diagnostic assays to detect and quantitate the presence of similar
 CC nucleic acids in samples, and therefore which patients may be in need of
 CC restorative therapy. The protein may also be used as antigens in the
 CC production of antibodies against neuregulin 1 and in assays to identify
 CC modulators of neuregulin 1 expression and activity. The antibodies and
 CC antagonists may also be used to down regulate expression and activity.
 CC The antibodies may also be used as diagnostic agents for detecting the
 CC presence of neuregulin 1 in samples.

SO Sequence 1503900 BP; 452487 A; 281874 C; 288074 G; 480092 T; 1373 other;

Query Match 61.3%; Score 18.4; DB 22; Length 1503900;
 Best Local Similarity 78.6%; Pred. No. 3.1e+02;
 Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CCTGTGCTAGTGTGACGACCTCTGCAG 28
 Db 801368 CCTGTGCTAGTGTGACGACGACCTCTGCAG 801341

RESULT 14
 ABX82153
 ID ABX82153 standard; cDNA; 288 BP.

XX ABX82153;

DT 24-APR-2003 (first entry)

DE Corn ear-derived polynucleotide (cpd) #613.

XX Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022;

XX SATMON023; structural gene; functional gene; regulatory gene;

XX corn ear-specific profile; gene transcription; gene expression;

XX hybrid plant; desirable trait expression; plant breeding program;

XX inheritance; desired characteristic; growth; development;

XX disease resistance; environmental adaptability; quality; yield;

XX multigene trait; plant; gene; ss.

XX Zea mays.

XX US6476212-B1.

XX 05-NOV-2002.

XX 14-MAY-1999; 99US-0313294.

XX 26-MAY-1998; 98US-086722P.

XX (INCY-) INCYTE GENOMICS INC.

XX Laljundi RV, Ito LY, Sherman BK;

XX WPI; 2003-208840/20.

XX Novel purified corn-ear derived polynucleotide useful as hybridization

PT probe for detecting polynucleotide in sample, and for identifying,

PT evaluating, and altering desired characteristics associated with

XX growth, development -

XX Examples; SEQ ID NO 613; 390pp; English.

CC The present invention relates to the isolation of corn ear-derived
 CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries
 CC SATMON022 and SATMON023. Some of the cpds uniquely identify structural,
 CC functional, and regulatory genes of corn ear. The polynucleotides
 CC sequences are useful for detecting cpds in a sample, for producing
 CC a corn ear-specific profile of gene transcription, for detecting
 CC altered gene expression in inbred or hybrid plants, and for screening
 CC several molecules for specific binding to the polynucleotide. The cpds
 CC are useful to identify, isolate, or extend identical or related
 CC corn-ear nucleic acid sequences from DNA libraries, and in nucleic

CC acid amplification or hybridisation techniques to follow the
 CC expression of desirable traits through plant breeding programs.
 CC Preferably, the cdps are used to identify, evaluate, alter, or
 CC follow the inheritance of desired characteristics associated with
 CC growth and development, disease resistance, environmental adaptability,
 CC quality, and yield of corn. The cdps are also useful as molecular
 CC markers for studying inheritance and multigene traits in a plant
 CC breeding program. The cdps are useful for producing purified corn-ear
 CC polypeptides by recombinant techniques. They are also useful in
 CC diagnostic assays to detect or confirm conditions or diseases
 CC associated with abnormal levels of cdp expression. ABX81541-ABX89140
 CC represent corn ear-derived polynucleotides (cdps) of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site at seqdata.uspto.gov/psipd/IDentry.html.
 CC
 SQ Sequence 288 BP, 55 A, 108 C, 81 G, 44 T, 0 other;
 Query Match 59.3%; Score 17.8; DB 25; Length 288;
 Best Local Similarity 90.5%; Pred. No. 3.4e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 10 GTGCTACGACCTCTGCGAGG 30
 DB 208 GTGCTACGACCTCTGCGAGG 228
 RESULT 15
 ABL72764/C
 ID ABL72764 standard; cDNA; 289 BP.
 AC ABL72764;
 XX
 DT 14-MAY-2002 (first entry)
 XX
 DE Corn tassels-derived polynucleotide (cdps) SEQ ID NO:2138.
 XX
 KW Corn; corn tassels-derived polynucleotide; cdps; hybrid breeding; CDPs;
 KW inheritance; characteristic; growth; development; disease resistance;
 KW environmental adaptability; quality; yield; molecular marker;
 KW multigene trait; plant breeding; corn tassels; gene; ss.
 XX
 OS Zea mays.
 XX
 PN US2001051335-A1.
 XX
 PD 13-DEC-2001.
 XX
 PF 16-APR-1999; 99US-0294093.
 XX
 PR 21-APR-1998; 98US-082567P.
 XX
 PA (LALG/) LALGUDI R V.
 PA (ITOL/) ITO L Y.
 PA (SHER/) SHERMAN B K.
 XX
 PI Lalgudi RV, Ito LY, Sherman BK;
 XX
 DR MPI; 2002-163647/21.
 XX
 PT Novel purified corn tassels-derived polynucleotide useful for
 PT determining altered gene expression, to recover regulatory elements and
 PT to follow inheritance of desirable characteristics through hybrid
 PT breeding programs -
 XX
 PS Claim 1; SEQ ID 2138; 201pp; English.
 XX
 CC The present sequence describes a purified corn tassels-derived
 CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
 CC selected from those given in ABL70627 to ABL76833. The cdps sequences
 CC encode corn tassels-derived polypeptides (CDPs). The cdps sequences (I)
 CC can be used for determining altered gene expression, to recover
 CC regulatory elements and to follow inheritance of desirable

CC characteristics through hybrid breeding programs. (I) are also useful
 CC in the evaluation, and alteration of desired characteristics associated
 CC with growth and development, disease resistance, environmental
 CC adaptability, quality and yield, and as molecular markers for studying
 CC inheritance of multigene traits in a plant breeding program. (I) can be
 CC used to produce a tassels-specific profile of gene transcription, a
 CC transcript image, to clone regulatory elements for use in transformation
 CC vectors, to express a polypeptide, to identify, isolate or extend
 CC identical or related corn tassels nucleic acid sequences from DNA
 CC libraries, in nucleic acid hybridisation or amplification technologies,
 CC as query sequences to determine homology of known sequences, as probe
 CC for use in Southern or Northern hybridisation, and to identify the
 CC presence of and/or to determine the degree of similarity between two
 CC (or more) nucleic acid sequences.
 CC
 SQ Sequence 289 BP, 79 A, 51 C, 75 G, 84 T, 0 other;
 Query Match 59.3%; Score 17.8; DB 24; Length 289;
 Best Local Similarity 75.9%; Pred. No. 3.4e+02;
 Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 2 CCTGCGTAGTGCTACGACCTCTGCGAGG 30
 DB 227 CCTATATAGTAGTATGACCACTGCGAGG 199

Search completed: September 16, 2003, 18:37:05
 Job time : 157.297 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using BW model

Run on: September 16, 2003, 18:24:35 ; Search time 38.3168 Seconds
(without alignments)
345.579 Million cell updates/sec

Title: US-09-594-065-3

Perfect score: 30
Sequence: 1 CCCGCGTAGTGTGACGACCTCCGACGAG 30

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19.4	64.7	483	US-09-252-991A-6704	Sequence 6704, Ap
C 2	19.4	64.7	510	US-09-252-991A-6599	Sequence 6599, Ap
C 3	19.4	64.7	3084	US-09-252-991A-6539	Sequence 6539, Ap
C 4	19.4	64.7	3273	US-09-252-991A-6578	Sequence 6578, Ap
C 5	18.8	62.7	1368	US-08-738-172-3	Sequence 3, Appli
C 6	18.8	62.7	2936	US-08-738-172-1	Sequence 1, Appli
C 7	18.8	62.7	4454	US-08-738-172-2	Sequence 2, Appli
C 8	17.8	59.3	288	US-09-313-294A-613	Sequence 613, App
C 9	17.8	59.3	1766	US-09-904-615-29	Sequence 29, Appli
C 10	17.8	59.3	2035	PCT-UB93-12464-1	Sequence 1, Appli
C 11	17.8	59.3	12848	US-09-453-702B-252	Sequence 252, App
C 12	17.8	59.3	48908	US-09-453-702B-137	Sequence 137, App
C 13	17.4	58.0	1520	US-09-620-312D-458	Sequence 458, App
C 14	17.2	57.3	8460	US-08-469-005A-9	Sequence 9, Appli
C 15	17.2	57.3	8519	US-09-261-907-1	Sequence 1, Appli
C 16	17.2	57.3	11725	US-08-756-506-1	Sequence 1, Appli
C 17	17.2	57.3	49377	US-08-764-233A-1	Sequence 14975, A
C 18	17.2	56.7	1803	US-09-252-991A-14975	Sequence 14842, A
C 19	17.2	56.7	2076	US-09-252-991A-14842	Sequence 913, App
C 20	16.8	56.0	4530	US-09-221-017B-913	Sequence 460, App
C 21	16.8	56.0	100	US-09-060-756-460	Sequence 460, App
C 22	16.8	56.0	100	US-09-670-314-460	Sequence 18, Appli
C 23	16.8	56.0	1214	US-09-506-286B-15	Sequence 15, Appli
C 24	16.8	56.0	1214	US-09-506-286B-18	Sequence 18, Appli
C 25	16.8	56.0	1214	US-09-762-861B-15	Sequence 15, Appli
C 26	16.8	56.0	1214	US-09-762-861B-18	Sequence 18, Appli
C 27	16.8	56.0	1230	US-09-252-991A-7646	Sequence 7646, Ap

28	16.8	56.0	1241	4	US-09-506-286B-13	Sequence 13, Appli
29	16.8	56.0	1241	4	US-09-506-286B-16	Sequence 16, Appli
30	16.8	56.0	1241	4	US-09-762-861B-13	Sequence 13, Appli
31	16.8	56.0	1241	4	US-09-762-861B-16	Sequence 16, Appli
C 32	16.8	56.0	1270	4	US-09-016-434-756	Sequence 756, App
C 33	16.8	56.0	1445	4	US-09-697-367-19	Sequence 19, Appli
C 34	16.8	56.0	2104	1	US-08-592-128-96	Sequence 96, Appli
35	16.8	56.0	2104	4	US-09-168-595-96	Sequence 96, Appli
36	16.8	56.0	2194	2	US-08-633-879C-1	Sequence 3, Appli
37	16.8	56.0	2277	4	US-09-506-286B-46	Sequence 46, Appli
38	16.8	56.0	2277	4	US-09-506-286B-49	Sequence 49, Appli
39	16.8	56.0	2341	4	US-09-506-286B-44	Sequence 44, Appli
40	16.8	56.0	2341	4	US-09-506-286B-47	Sequence 47, Appli
41	16.8	56.0	1230025	4	US-09-198-452A-1	Sequence 1, Appli
C 42	16.4	54.7	432	4	US-09-397-787-191	Sequence 191, Appli
43	16.4	54.7	447	4	US-09-183-861-78	Sequence 78, Appli
44	16.4	54.7	447	4	US-09-022-765-78	Sequence 78, Appli
45	16.4	54.7	447	4	US-09-551-974A-78	Sequence 78, Appli

ALIGNMENTS

```
RESULT 1
US-09-252-991A-6704/C
Sequence 6704, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubinfeld et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 6704
LENGTH: 483
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6704

Query Match          64.7% Score 19.4; DB 4; Length 483;
Best Local Similarity 79.3%; Pred. No. 10;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCCGCGTAGTGTGACGACCTCCGACGAG 29
DB 35 CCCGCGTAGTGTGACGACCTCCGACGAG 7

RESULT 2
US-09-252-991A-6599
Sequence 6599, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubinfeld et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 6599
LENGTH: 510
TYPE: DNA
```

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6599

Query Match 64.7%; Score 19.4; DB 4; Length 510;
Best Local Similarity 79.3%; Pred. No. 10;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 CCTGCGTAGTGTAGTACGACCTCTCGAGG 29
|||||
Db 347 CCCGCGCAGTGGCGGAGCGCTCTGAGG 375

RESULT 3
US-09-252-991A-6639/c
Sequence 6639, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 6639
LENGTH: 3084
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6639

Query Match 64.7%; Score 19.4; DB 4; Length 3084;
Best Local Similarity 79.3%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 CCTGCGTAGTGTAGTACGACCTCTCGAGG 29
|||||
Db 1309 CCCGCGCAGTGGCGGAGCGCTCTGAGG 1281

RESULT 4
US-09-252-991A-6578
Sequence 6578, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 6578
LENGTH: 3273
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6578

Query Match 64.7%; Score 19.4; DB 4; Length 3273;
Best Local Similarity 79.3%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 CCTGCGTAGTGTAGTACGACCTCTCGAGG 29
|||||
Db 1644 CCCGCGCAGTGGCGGAGCGCTCTGAGG 1672

RESULT 5
US-08-738-172-3
Sequence 3, Application US/08738172
Patent No. 5939257

GENERAL INFORMATION:
APPLICANT: Szasz, Joseph
TITLE OF INVENTION: THERMOSTABLE ALKALINE PHOSPHATASES
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/738,172
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/005,965
FILING DATE: October 27, 1995
APPLICATION NUMBER: 08/465,003
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/575,354
FILING DATE: December 20, 1995
APPLICATION NUMBER: 08/240,158
FILING DATE: May 10, 1994
APPLICATION NUMBER: 08/229,329
FILING DATE: April 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 222/158
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1368 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-738-172-3

Query Match 62.7%; Score 18.8; DB 2; Length 1368;
Best Local Similarity 76.7%; Pred. No. 21;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

1 CCTGCGTAGTGTAGTACGACCTCTCGAGG 30
|||||
Db 197 CCGTGGTACTGTCTCGGCTCCAGCCGGG 226

RESULT 6
US-08-738-172-1/c
Sequence 1, Application US/08738172
Patent No. 5939257

GENERAL INFORMATION:
APPLICANT: Szasz, Joseph
TITLE OF INVENTION: THERMOSTABLE ALKALINE PHOSPHATASES
NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/738,172
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/005,965
FILING DATE: October 27, 1995
APPLICATION NUMBER: 08/465,003
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/575,354
FILING DATE: December 20, 1995
APPLICATION NUMBER: 08/240,158
FILING DATE: May 10, 1994
APPLICATION NUMBER: 08/229,329
FILING DATE: April 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 222/158
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2936 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-738-172-1

Query Match 62.7%; Score 18.8; DB 2; Length 2936;
Best Local Similarity 76.7%; Pred. No. 23;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCTGCGTAGTGTGACGACCTCTCCAGGG 30
DB 1351 CCGTGGTACTGCTTCGGGCTCCAGCCGGG 1322

RESULT 7
US-08-738-172-2/c
Sequence 2, Application US/08738172
Patent No. 5939257
GENERAL INFORMATION:
APPLICANT: Szasz, Joseph
APPLICANT: Davis, Maria
TITLE OF INVENTION: THERMOSTABLE ALKALINE PHOSPHATASES
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/738,172
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/005,965
FILING DATE: October 27, 1995
APPLICATION NUMBER: 08/465,003
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/575,354
FILING DATE: December 20, 1995
APPLICATION NUMBER: 08/240,158
FILING DATE: May 10, 1994
APPLICATION NUMBER: 08/229,329
FILING DATE: April 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 222/158
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4454 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-738-172-2

Query Match 62.7%; Score 18.8; DB 2; Length 4454;
Best Local Similarity 76.7%; Pred. No. 24;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCTGCGTAGTGTGACGACCTCTCCAGGG 30
DB 1738 CCGTGGTACTGCTTCGGGCTCCAGCCGGG 1709

RESULT 8
US-09-313-294A-613
Sequence 613, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Ialagudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 613
LENGTH: 288
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700549530H1
US-09-313-294A-613

Query Match 59.3%; Score 17.8; DB 4; Length 288;
Best Local Similarity 90.5%; Pred. No. 48;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 GTGGTACGACCTCTCCAGGG 30
|||||

Db 208 GTGTACGACTCCGGCTGAG 228

RESULT 9

US-09-904-615-29/c
Sequence 29, Application US/09904615
Patent No. 6566325
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 49 Human Secreted Proteins
FILE REFERENCE: P2032P1
CURRENT APPLICATION NUMBER: US/09/904,615
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 09/511,554
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/097,917
PRIOR FILING DATE: 1998-08-25
PRIOR APPLICATION NUMBER: 60/098,634
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 170
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 29
LENGTH: 1766
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1743)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1748)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1749)
OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-29

Query Match 59.3%; Score 17.8; DB 4; Length 1766;
Best Local Similarity 75.9%; Pred. No. 59;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CCTCGGTAGTGTACGACTCTCTGAG 29
Db 1312 CCTGGGTGTGCTTCCACTCTGAG 1284

RESULT 10

PCT-US93-12464-1
Sequence 1, Application PC/TUS9312464
GENERAL INFORMATION:
APPLICANT: The Regents of the University of California, et al.
TITLE OF INVENTION: METHOD FOR INHIBITION OF CELL
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPENSLER HORN JUBAS & LUBITZ
STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
CITY: LOS ANGELES
STATE: CALIFORNIA
COUNTRY: US
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12464
FILING DATE: 21-DEC-1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: HOWELLS, STACY L.
REGISTRATION NUMBER: 34,842

REFERENCE/DOCKET NUMBER: PD2929
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 1,
SEQUENCE CHARACTERISTICS:
LENGTH: 2035 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: ALPHA-L-FUCOSIDASE
FEATURE:
NAME/KEY: CDS
LOCATION: 19..1401
PCT-US93-12464-1

Query Match 59.3%; Score 17.8; DB 5; Length 2035;
Best Local Similarity 75.9%; Pred. No. 60;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CCTCGGTAGTGTACGACTCTCTGAG 29
Db 351 CCGGAGGAGTGGCCGACTCTTCCAG 379

RESULT 11

US-09-453-702B-252/c
Sequence 252, Application US/09453702B
Patent No. 6365723
GENERAL INFORMATION:
APPLICANT: Blactner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Plunkney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296,95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 252:
SEQUENCE CHARACTERISTICS:
LENGTH: 12848
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 252:

US-09-453-702B-252

Query Match 59.3%; Score 17.8; DB 4; Length 12848;
Best Local Similarity 75.9%; Pred. No. 74;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCTGCGTAGGTAGTACCTCTCGACGG 30
DB 4539 CTTGCTGAGCGGAGACCTCGTGCAGG 4511

RESULT 12

US-09-453-702B-137/c

Sequence 137, Application US/09453702B

Patent No. 6365723

GENERAL INFORMATION:

APPLICANT: Blatner, Frederick R.

Burland, Nicole T.

Punkett, Guy

Welch, Rod

TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSER: Quarles & Brady

STREET: 1 South Pinckney Street

CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53701-2113

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/453,702B

FILING DATE: 03-Dec-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/110,955

FILING DATE: 04-Dec-1998

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 960296,95017

TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 251-5000

INFORMATION FOR SEQ ID NO: 137:

SEQUENCE CHARACTERISTICS:

LENGTH: 48908

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 137:

US-09-453-702B-137

Query Match 59.3%; Score 17.8; DB 4; Length 48908;
Best Local Similarity 75.9%; Pred. No. 86;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCTGCGTAGGTAGTACCTCTCGACGG 30
DB 15720 CTTGCTGAGCGGAGACCTCGTGCAGG 15692

RESULT 13

US-09-620-312D-458

Sequence 458, Application US/09620312D

Patent No. 6569862

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungling
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Dymnac, Radoje T.
TITLE OF INVENTION: No. 656962el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pf_fl_genes Version 1.0
SEQ ID NO 458
LENGTH: 1520
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (271)..(1161)
US-09-620-312D-458

Query Match 58.0%; Score 17.4; DB 4; Length 1520;
Best Local Similarity 77.8%; Pred. No. 87;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCTGCGTAGGTAGTACGACCTCTGCGAG 28
DB 277 CCGGAGTCTGTGACGACCTGTCGAG 303

RESULT 14

US-08-469-005A-9/c

Sequence 9, Application US/08469005A

Patent No. 565874

GENERAL INFORMATION:

APPLICANT: KUHADJA, FRANCIS P.

APPLICANT: PASTERNAK, GARY A.

TITLE OF INVENTION: CANCER RELATED ANTIGEN

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: BAKER & BOTT, L.L.P.

STREET: 1299 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20004-2400

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,005A

FILING DATE: 05-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/188,426

FILING DATE: 24-JAN-1994
 APPLICATION NUMBER: 08/096,908
 FILING DATE: 26-JUL-1993
 APPLICATION NUMBER: 07/917,716
 FILING DATE: 24-JUL-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Posortake, Laurence H
 REGISTRATION NUMBER: 34,698
 REFERENCE/DOCKET NUMBER: 062482-0113
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-639-7700
 TELEFAX: 202-639-7890
 TELEX:
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8460 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE:
 ORIGINAL SOURCE:
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 124...7650
 OTHER INFORMATION:
 US-08-469-005A-9

Query Match 57.3%; Score 17.2; DB 1; Length 8460;
 Best Local Similarity 73.3%; Pred. No. 1.3e+02;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCTGCGTAGTGCTAGCACTCCTCGAGG 30
 DB 3550 CCTGCGACAGTGCAGCTCCTCCTCGAGG 3521

RESULT 15
 US-09-261-907-1/C
 Sequence 1, Application US/09261907A
 Patent No. 6294364
 GENERAL INFORMATION:
 APPLICANT: ELLIS, CATHERINE
 APPLICANT: LONSDALE, JOHN
 APPLICANT: BERGMA, DEBK J.
 APPLICANT: MOONEY, JEFFREY L.
 APPLICANT: DEPIERA, MEGAN E.
 APPLICANT: CHAPMAN, CONRAD
 TITLE OF INVENTION: HUMAN PAS
 FILE REFERENCE: GP-70603
 CURRENT APPLICATION NUMBER: US/09/261,907A
 CURRENT FILING DATE: 1999-03-03
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 1
 LENGTH: 8519
 TYPE: DNA
 ORGANISM: HOMO SAPIENS
 US-09-261-907-1

Query Match 57.3%; Score 17.2; DB 3; Length 8519;
 Best Local Similarity 73.3%; Pred. No. 1.3e+02;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCTGCGTAGTGCTAGCACTCCTCGAGG 30
 DB 3582 CCTGCGACAGTGCAGCTCCTCCTCGAGG 3553

Search completed: September 16, 2003, 20:42:38
 Job time : 40.3168 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 19:14:25 ; Search time 108.416 Seconds
(without alignments)
680.690 Million cell updates/sec

Title: US-09-594-065-3

Perfect score: 30
Sequence: 1 cccgcgtagtgcgtacgacctccgcaggg 30

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1660708 seqs, 1229959015 residues

Total number of hits satisfying chosen parameters: 3321416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/FCI_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/FCIS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19.4	64.7	1323	14 US-10-156-761-4750	Sequence 4750, App1
2	19.4	64.7	110079	14 US-10-175-523-96	Sequence 96, App1
3	19.4	64.7	9025608	14 US-10-156-761-1	Sequence 1, App1
4	18.8	62.7	576	10 US-09-738-626-2147	Sequence 2147, App
5	18.8	62.7	1368	9 US-09-191-989-3	Sequence 3, App1
6	18.8	62.7	1593	14 US-10-156-761-5071	Sequence 5071, App
7	18.8	62.7	2936	9 US-09-191-989-1	Sequence 1, App1
8	18.8	62.7	4434	9 US-09-191-989-2	Sequence 2, App1
9	18.8	62.7	99916	10 US-09-816-095-3	Sequence 3, App1
10	18.8	62.7	3309400	10 US-09-738-626-1	Sequence 39424, App1
11	18.4	61.3	536	13 US-10-027-632-39424	Sequence 39424, A
12	18.4	61.3	536	13 US-10-027-632-58202	Sequence 58202, A
13	18.4	61.3	536	13 US-10-027-632-58209	Sequence 58209, A
14	18.4	61.3	536	13 US-10-027-632-309173	Sequence 309173, A
15	18.4	61.3	4705	12 US-10-312-918-1	Sequence 1, App1
16	18.4	61.3	4786	12 US-10-312-918-3	Sequence 3, App1

C 17	18.4	61.3	1503841	9 US-09-795-668-1	Sequence 1, App1
C 18	18.4	61.3	1503841	9 US-09-795-668-1	Sequence 1, App1
C 19	18.4	61.3	1503841	10 US-09-946-807-1	Sequence 1, App1
C 20	17.8	59.3	289	9 US-09-294-093B-2138	Sequence 2138, App
C 21	17.8	59.3	630	14 US-10-156-761-7444	Sequence 7444, App
C 22	17.8	59.3	745	13 US-10-027-632-146355	Sequence 146355, App
C 23	17.8	59.3	1766	9 US-09-738-254-29	Sequence 29, App1
C 24	17.8	59.3	1766	9 US-09-904-615-29	Sequence 29, App1
C 25	17.8	59.3	1766	12 US-10-055-998-29	Sequence 29, App1
C 26	17.8	59.3	1766	14 US-10-054-988-29	Sequence 29, App1
C 27	17.8	59.3	2035	12 US-09-873-367C-711	Sequence 711, App
C 28	17.8	59.3	2052	14 US-10-156-761-4847	Sequence 4847, App
C 29	17.8	59.3	2162	14 US-10-158-846-13433	Sequence 13433, A
C 30	17.8	59.3	12848	12 US-10-205-194-30	Sequence 30, App1
C 31	17.8	59.3	12848	14 US-10-114-170-252	Sequence 252, App
C 32	17.8	59.3	48908	14 US-10-114-170-137	Sequence 137, App
C 33	17.8	59.3	9025608	14 US-10-156-761-1	Sequence 1, App1
C 34	17.6	58.7	4829	12 US-10-021-660-36	Sequence 36, App
C 35	17.6	58.7	12404	13 US-10-108-605-236	Sequence 236, App
C 36	17.4	58.0	81	10 US-09-783-590-7470	Sequence 7470, App
C 37	17.4	58.0	403	11 US-09-918-995-16763	Sequence 16763, A
C 38	17.4	58.0	606	10 US-09-951-470-6	Sequence 6, App1
C 39	17.4	58.0	700	13 US-10-027-632-150978	Sequence 150978, App
C 40	17.4	58.0	700	13 US-10-027-632-150979	Sequence 150979, App
C 41	17.4	58.0	700	13 US-10-027-632-150980	Sequence 150980, App
C 42	17.4	58.0	799	9 US-09-764-870-263	Sequence 263, App
C 43	17.4	58.0	799	14 US-10-125-540-863	Sequence 263, App
C 44	17.4	58.0	879	10 US-09-951-470-4	Sequence 4, App1
C 45	17.4	58.0	1520	14 US-10-037-270-458	Sequence 458, App

ALIGNMENTS

RESULT 1
US-10-156-761-4750
Sequence 4750, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIORITY APPLICATION NUMBER: JP 2001-204089
PRIORITY FILING DATE: 2001-05-30
PRIORITY APPLICATION NUMBER: JP 2001-272697
PRIORITY FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 4750
LENGTH: 1323
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (1323)
US-10-156-761-4750

Query Match 64.7%; Score 19.4; DB 14; Length 1323;
Best Local Similarity 79.3%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCCGCGTAGTGCgtacgacctccgcaggg 29
DB 669 CCCGCGTAGTGCgtacgacctccgcaggg 697

```
RESULT 2
US-10-175-523-96/c
; Sequence 96, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Paleyman, Michael
; APPLICANT: Rajan, Priti
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPTS)
; FILE REFERENCE: 3235/11795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 96
; LENGTH: 110079
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(110079)
; OTHER INFORMATION: where n may be a or g or c or t/u, unknown, or other
US-10-175-523-96

Query Match      64.7%; Score 19.4; DB 14; Length 110079;
Best Local Similarity 79.3%; Pred. No. 26;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1  CCCTGCGTAGTGTGATGACGACCTCTCGAGG 29
Db      4399  CACTGCTAGTGTGATGATCTCTCGAGG 4371

RESULT 3
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
```

```
; ORGANISM: Streptomyces avermectilis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match      64.7%; Score 19.4; DB 14; Length 9025608;
Best Local Similarity 79.3%; Pred. No. 23;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1  CCCTGCGTAGTGTGATGACGACCTCTCGAGG 29
Db      5809596  CCCCCGAAAGCGGTGACGACGTCCTCTCG 5809568

RESULT 4
US-09-738-626-2147/c
; Sequence 2147, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 2147
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-2147

Query Match      62.7%; Score 18.8; DB 10; Length 576;
Best Local Similarity 90.9%; Pred. No. 51;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      9  AGTGTACGACCTCTCGAGG 30
Db      258  AGTGTACGACCACTCGAGG 237

RESULT 5
US-09-191-989-3
; Sequence 3, Application US/09191989
; Patent No. US2002005098A1
; GENERAL INFORMATION:
; APPLICANT: Szaez, Joseph
; APPLICANT: Davis, Maria
; TITLE OF INVENTION: THERMOSTABLE ALKALINE PHOSPHATASES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: California
```

```

; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191.989
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/738,172
; FILING DATE: October 25, 1996
; APPLICATION NUMBER: 08/465,003
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/575,354
; FILING DATE: December 20, 1995
; APPLICATION NUMBER: 08/240,158
; FILING DATE: May 10, 1994
; APPLICATION NUMBER: 08/229,329
; FILING DATE: April 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wardburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 222/158
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1368 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-191-989-3

Query Match          62.7%; Score 18.8; DB 9; Length 1368;
Best Local Similarity 76.7%; Pred. No. 50;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1  CCTGCGTAGTGATGACGACTCTGCGAGGG 30
Db      197  CGGTGCGTAGTGCTTCCGGCTTCCAGCGGG 226

RESULT 6
US-10-156-761-5071
; Sequence 5071, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 5071
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1593)
; US-10-156-761-5071

Query Match          62.7%; Score 18.8; DB 14; Length 1593;
Best Local Similarity 76.7%; Pred. No. 50;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1  CCTGCGTAGTGATGACGACTCTGCGAGGG 30
Db      720  CGCGCGCAGCGGTACGACATCTGCTGGG 749

RESULT 7
US-09-191-989-1/c
; Sequence 1, Application US/09191989
; Patent No. US20020055098A1
; GENERAL INFORMATION:
; APPLICANT: Szaez, Joseph
; APPLICANT: Davis, Maria
; TITLE OF INVENTION: THERMOSTABLE ALKALINE PHOSPHATASES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191.989
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/738,172
; FILING DATE: October 25, 1996
; APPLICATION NUMBER: 08/465,003
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/575,354
; FILING DATE: December 20, 1995
; APPLICATION NUMBER: 08/240,158
; FILING DATE: May 10, 1994
; APPLICATION NUMBER: 08/229,329
; FILING DATE: April 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wardburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 222/158
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2936 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-191-989-1

Query Match          62.7%; Score 18.8; DB 9; Length 2936;
Best Local Similarity 76.7%; Pred. No. 50;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1  CCTGCGTAGTGATGACGACTCTGCGAGGG 30
```

Db 1351 CCCTGCGTACTGCTCGCCTCCAGCCGGG 1322

RESULT 8

US-09-191-989-2/c
; Sequence 2, Application US/09191989
; Patent No. US20020055098A1
; GENERAL INFORMATION:
; APPLICANT: Szasz, Joseph
; APPLICANT: Davis, Maria
; TITLE OF INVENTION: THERMOSTABLE ALKALINE PHOSPHATASES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09191,989
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/738,172
; FILING DATE: October 25,1996
; APPLICATION NUMBER: 08/465,003
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/575,354
; FILING DATE: December 20, 1995
; APPLICATION NUMBER: 08/240,158
; FILING DATE: May 10, 1994
; APPLICATION NUMBER: 08/229,329
; FILING DATE: April 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 222/158
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 955-0440
; TELEFAX: (213) 955-0440
; TELEX: 673510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4454 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-191-989-2

Query Match 62.7%; Score 18.8; DB 9; Length 4454;
Best Local Similarity 76.7%; Pred. No. 49;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCCTGCGTACTGCTCGCCTCCAGCCGGG 30
DB 1738 CCCTGCGTACTGCTCGCCTCCAGCCGGG 1709

RESULT 9

US-09-816-095-3/c
; Sequence 3, Application US/09816095
; Patent No. US20020137164A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weihua

; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CLO01147
; CURRENT APPLICATION NUMBER: US/09/816,095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 99916
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(99916)
; OTHER INFORMATION: n = A,T,C'or G
; US-09-816-095-3

Query Match 62.7%; Score 18.8; DB 10; Length 99916;
Best Local Similarity 76.7%; Pred. No. 48;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCCTGCGTACTGCTCGCCTCCAGCCGGG 30
DB 1748 CCCTGCGGCGCTGCGACGACCTCTCGAGGG 1719

RESULT 10

US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIRO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentin ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-1

Query Match 62.7%; Score 18.8; DB 10; Length 3309400;
Best Local Similarity 90.9%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 AGTGTAGCAGCTCTCTGCAAGG 30
DB 2066823 AGTGTAGCAGCAGCAGCTCGAAGG 2066802

RESULT 11

US-10-027-632-39424/c
; Sequence 39424, Application US/10027632
; GENERAL INFORMATION:

```
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39424
LENGTH: 536
TYPE: DNA
ORGANISM: Human
US-10-027-632-39424
```

```
Query Match          61.3%; Score 18.4; DB 13; Length 536;
Best Local Similarity 78.6%; Pred. No. 76;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 3 CTGCGTAGTGTGACGACCTCTCCGAGGG 30
    |||||  ||  |||||  |||||
Db 117 CTGCGTAGTGTGACGACCTCTCCGAGGG 90
```

```
RESULT 12
US-10-027-632-58202/c
Sequence 58202, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 58202
LENGTH: 536
TYPE: DNA
ORGANISM: Human
US-10-027-632-58202
```

```
Query Match          61.3%; Score 18.4; DB 13; Length 536;
Best Local Similarity 78.6%; Pred. No. 76;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 3 CTGCGTAGTGTGACGACCTCTCCGAGGG 30
    |||||  ||  |||||  |||||
Db 117 CTGCGTAGTGTGACGACCTCTCCGAGGG 90
```

```
RESULT 13
US-10-027-632-59209/c
Sequence 59209, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 59209
LENGTH: 536
TYPE: DNA
ORGANISM: Human
US-10-027-632-59209
```

```
Query Match          61.3%; Score 18.4; DB 13; Length 536;
Best Local Similarity 78.6%; Pred. No. 76;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 3 CTGCGTAGTGTGACGACCTCTCCGAGGG 30
    |||||  ||  |||||  |||||
Db 117 CTGCGTAGTGTGACGACCTCTCCGAGGG 90
```

```
RESULT 14
US-10-027-632-309173/c
Sequence 309173, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 309173
```

/ LENGTH: 536
 / TYPE: DNA
 / ORGANISM: Human
 US-10-027-632-309173

Query Match 61.3%; Score 18.4; DB 13; Length 536;
 Best Local Similarity 78.6%; Pred. No. 76;
 Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CTGCGTAGTGATGACGACCTCTGCGAGG 30
 ||||| ||||| ||||| ||||| |||||
 DB 117 CTGCGTAGTGACGACCTCTGCGAGG 90

RESULT 15

US-10-312-918-1
 / Sequence 1, Application US/10312918
 / Publication No. US20030162740A1
 / GENERAL INFORMATION:
 / APPLICANT: Regents of the University of California
 / APPLICANT: Wang, Jean Y.J.
 / APPLICANT: Vigneri, Paolo
 / TITLE OF INVENTION: New Strategy for Leukemia Therapy
 / FILE REFERENCE: 6627-PC0159
 / CURRENT APPLICATION NUMBER: US/10/312,918
 / CURRENT FILING DATE: 2002-12-27
 / PRIOR APPLICATION NUMBER: 60/215,595
 / PRIOR FILING DATE: 2000-06-30
 / NUMBER OF SEQ ID NOS: 4
 / SOFTWARE: Patent version 3.1
 / SEQ ID NO 1
 / LENGTH: 4705
 / TYPE: DNA
 / ORGANISM: Homo sapien/ Mus musculus
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (1)..(4704)
 / OTHER INFORMATION:
 US-10-312-918-1

Query Match 61.3%; Score 18.4; DB 12; Length 4705;
 Best Local Similarity 78.6%; Pred. No. 75;
 Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCTGCGTAGTGATGACGACCTCTGCGAGG 29
 ||||| ||||| ||||| ||||| |||||
 DB 3414 CCTGCGTAGTGATGACGACCTCTGCGAGG 3441

Search completed: September 16, 2003, 22:54:38
 Job time : 131.416 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 16, 2003, 18:22:25 / Search time 1518.42 Seconds
(without alignments)
480.194 Million cell updates/sec

Title: US-09-594-065-3

Perfect score: 30

Sequence: 1 CCCGCGTAGTGTGACGACCTCCTGCAAGG 30

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pln:*
20: em_ges_vit:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rod:*
26: em_ges_phg:*
27: em_ges_vrl:*
28: gb_gesi:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	22.2	74.0	900	10	BG288880 60238382
2	21	70.0	993	14	BY707129 BY707129
3	20.4	66.0	337	13	BX264906 BX264906
4	20.4	66.0	516	12	BM864032 mgcm008x1

Result No.	Score	Match	Length	ID	Description
5	20.4	68.0	880	29	CNS04LYM
6	20	66.7	752	9	AA896672
7	20	66.7	723	12	B1327791
8	19.8	66.0	450	11	BY463661
9	19.6	65.3	161	14	CD475739
10	19.4	64.7	380	10	BB872020
11	19.4	64.7	406	9	AA709478
12	19.4	64.7	432	13	BY680656
13	19.4	64.7	478	14	CB790399
14	19.4	64.7	434	10	BB857961
15	19.4	64.7	531	10	BF022130
16	19.4	64.7	576	12	B1888644
17	19.4	64.7	623	28	A2634235
18	19.4	64.7	633	10	BE269697
19	19.4	64.7	720	10	BG298602
20	19.4	64.7	724	12	B1692432
21	19.4	64.7	884	14	CA977558
22	19.4	64.7	937	10	BF106954
23	19.4	64.7	997	29	CNS040LD
24	19.4	64.7	1160	29	CC254336
25	19.4	64.7	2225	11	AK007059
26	19.4	64.7	2609	11	BC005727
27	19.4	64.7	2869	11	AK042427
28	19.2	64.0	380	14	CA690063
29	19.2	64.0	402	14	CA666616
30	19.2	64.0	555	14	CA666652
31	19.2	64.0	584	14	CA666706
32	19.2	64.0	241	10	BB261816
33	19.2	64.0	432	28	AQ216405
34	19.2	64.0	568	9	AA69480
35	19.2	64.0	623	10	BF252994
36	19.2	64.0	629	10	BE305546
37	19.2	64.0	794	29	CNS03DKF
38	19.2	64.0	814	29	CNS03OAP
39	19.2	64.0	840	29	CNS04FIP
40	19.2	64.0	1105	12	BM806374
41	18.8	62.7	300	9	AL838394
42	18.8	62.7	340	13	BY335529
43	18.8	62.7	366	2	HSW090445
44	18.8	62.7	391	14	BY702686
45	18.8	62.7	476	14	CB949437

ALIGNMENTS

RESULT 1
LOCUS BG288880
DEFINITION 602383822F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4512917 5', mRNA sequence.
ACCESSION BG288880
VERSION BG288880.1 GI:13044162
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 900)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contract: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLI at:
http://image.llnl.gov
Plate: LLM10398 row: 1 column: 06
High quality sequence stop: 718.

DEFINITION	BX264906 AGENAE Gallus gallus multi-tissues normalized and one-subtracted cDNA library (gcal)
VERSION	BX264906
KEYWORDS	BX264906.1 GI:28587504
SOURCE	EST.
ORGANISM	Gallus gallus (chicken)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
AUTHORS	1 (bases 1 to 337)
TITLE	Herault, F., Le Moutu-Metzing, V., Desert, C., Retout, E., Plum, F., Klopp, C. and Douaire, M.
JOURNAL	Construction and primary characterization of chicken normalized multi-tissue cDNA libraries
COMMENT	Unpublished
FEATURES	source
source	INRA, UMR INRA-ENSAR Genetique Animale 65 rue de Saint-Brieuc, Rennes cedex, 35042, FRANCE Tel: +33 (0) 2.23.48.54.63 Fax: +33 (0) 2.23.48.54.70 Email: Madeleine.Douaire@cazon.inra.fr Sequence cleaned of vector, adaptor and repetitions. Contact us at signasupport@jouy.inra.fr to obtain the chromatogram of this sequence. Plate: 0015 row: g column: 13 Seq primer: M13P.
Location/Qualifiers	1..337 /organism="Gallus gallus" /mol_type="mRNA" /db_xref="taxon:9031" /clone="gcal0015c.g.13" /tissue_type="multi-tissues" /dev_stage="from embryos to adults" /lab_host="DH10B" /clone_lib="AGENAE Gallus gallus multi-tissues normalized and one-subtracted cDNA library (gcal)" /note="Vector: pTT3D-pac; tissues: adipose tissue, brain, kidney, liver, multi-tissues, muscle, ovary, testis, bone marrow, caecum, duodenum, embryos, fabricius gland, granulosa, hypothalamus, ileon, jejunum, oviduct, pancreas, skin, spleen, thymus, utero-vaginal gland, pituitary gland, hematopoietic progenitor cells, small follicle. Clone distribution: AGENAE Resource centre, Francois Pium, Francois Plum, Inra.fr, INRA, CEA Radiobiologie et Etude du genome (UREG), Dominique de Vilvert, 78352, Jouy-en-Josas cedex, FRANCE"
BASE COUNT	73 a 108 c 77 g 79 t
ORIGIN	
Query Match	68.0% Score 20.4; DB 13; Length 337;
Best Local Similarity	80.0% Pred. No. 9.4e-02;
Matches	24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy	1 CCTGCTGATGCTAGCAGACTCTCTGACAGG 30
Dd	81 CCGGCCAAGGGGCAACGACTCTCTCAAGG 110
RESULT 4	
LOCUS	BM664032
DEFINITION	mgcm008x16.f.b Magnaporthe grisea CM uni-zap XR Library Magnaporthe
ACCESSION	BM664032
VERSION	BM664032.2 GI:30391021
KEYWORDS	EST.
SOURCE	Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes; Incertae sedis; Magnaportheaceae; Magnaporthe.

```

REFERENCE 1 (bases 1 to 516)
AUTHORS Ebbole,D.J., Yuan,J., Thomas,T.L., Bobrowicz,P., Lu,G., Bhatterai
TITLE 'K. and Dean,R.A.
GRISEA expressed sequence tags from the rice blast fungus, Magnaporthe
grisea
JOURNAL Unpublished
COMMENT On Mar 7, 2002 this sequence version replaced gi:19231714.

JOURNAL Unpublished
COMMENT On Mar 7, 2002 this sequence version replaced gi:19231714.
Contact: Ebbole DJ
Department of Plant Pathology & Microbiology
Texas A&M University
Peterson Bldg, MS2132, College Station, TX 77843-2132, USA
Tel: 979 845 4831
Fax: 979 845 4831
Email: d.ebbole@tamu.edu
Chromatogram file of this sequence is available, see contact person
/best nr hit (April 22, 2003) sp|P15710|PHO4_NECR
PHOSPHATE-RESISTIBLE PHOSPHATE PERMEASE >g... 53 2e-06
PCR primers
FORWARD: T3 primer
BACKWARD: T7 primer
Plate: mgcm008 row: I column: 16
Seq primer: T3.
Location/Qualifiers
1. 516
/organism="Magnaporthe grisea"
/mol_type="mRNA"
/strain="Guy11"
/db_xref="taxon:148305"
/clone="mgcm008x16"
/sex="Mat-2 hermaphrodite"
/cell_type="mycelium"
/clone_id="Magnaporthe grisea Cm Uni-2ap XR library"
/notes="Vector: pBluescriptSK-; Site 1: EcoRI; Site 2: XhoI
; Unidirectional cloning. EcoRI side has T3 primer and
predominantly 5' reads. T7 primer on XhoI side of insert.
Strain inoculated into complete medium grown for two days
at room temperature, 150 rpm, harvested, blended,
reincubated into complete medium 24 h, room temperature,
150 rpm. Sequences were processed by one of two methods.
Where a full-length alignment to the M. grisea genome
sequence was available, the EST sequence was trimmed
according to the alignment, otherwise sequence quality was
assessed using phredphrap version 991019 and trimmed
according to phd files (0.05) and for vector segs."
BASE COUNT 97 a 193 c 147 g 79 t
ORIGIN
Query Match 68.0%; Score 20.4; DB 12; Length 516;
Best Local Similarity 80.0%; Pred. No. 9.9e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

oy 1 CCTCGCGTAGGCGACGACCTCGCAGGG 30
||| ||||| ||||| ||||| |||||
Db 121 CACTCTGAAGTGGTACACATCTTCCTCAGGG 150

RESULT 5
CNS04LYM/c CNS04LYM 880 bp DNA linear GSS 01-SEP-2000
LOCUS Tetradon nigroviridis genome survey sequence PUC-ORI end of clone
DEFINITION 119008 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL296743.1 GI:8035323
VERSION AL296743
KEYWORDS GSS: genome survey sequence.
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetradon.
REFERENCE 1
Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,

```

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS
TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS
1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30
31	32	33	34	35	36
37	38	39	40	41	42
43	44	45	46	47	48
49	50	51	52	53	54
55	56	57	58	59	60
61	62	63	64	65	66
67	68	69	70	71	72
73	74	75	76	77	78
79	80	81	82	83	84
85	86	87	88	89	90
91	92	93	94	95	96
97	98	99	100	101	102
103	104	105	106	107	108
109	110	111	112	113	114
115	116	117	118	119	120
121	122	123	124	125	126
127	128	129	130	131	132
133	134	135	136	137	138
139	140	141	142	143	144
145	146	147	148	149	150
151	152	153	154	155	156
157	158	159	160	161	162
163	164	165	166	167	168
169	170	171	172	173	174
175	176	177	178	179	180
181	182	183	184	185	186
187	188	189	190	191	192
193	194	195	196	197	198
199	200	201	202	203	204
205	206	207	208	209	210
211	212	213	214	215	216
217	218	219	220	221	222
223	224	225	226	227	228
229	230	231	232	233	234
235	236	237	238	239	240
241	242	243	244	245	246
247	248	249	250	251	252
253	254	255	256	257	258
259	260	261	262	263	264
265	266	267	268	269	270
271	272	273	274	275	276
277	278	279	280	281	282
283	284	285	286	287	288
289	290	291	292	293	294
295	296	297	298	299	300
301	302	303	304	305	306
307	308	309	310	311	312
313	314	315	316	317	318
319	320	321	322	323	324
325	326	327	328	329	330
331	332	333	334	335	336
337	338	339	340	341	342
343	344	345	346	347	348
349	350	351	352	353	354
355	356	357	358	359	360
361	362	363	364	365	366
367	368	369	370	371	372
373	374	375	376	377	378
379	380	381	382	383	384
385	386	387	388	389	390
391	392	393	394	395	396
397	398	399	400	401	402

Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:670821
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..652
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:1279021"
/tissue_type="macrophage"
/dev_stage="MEH-3 cell line"
/lab_host="SOLR (kanamycin resistant)"
/clone_id="Stragene mouse macrophage (#937306)"
/note="Organ: blood; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. MEH-3 cell line. Average insert size: 1.5 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCGCGAG
3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'."

BASE COUNT
136 a 191 c 176 g 148 t 1 others

ORIGIN

Query Match 66.7%; Score 20; DB 9; Length 652;
Best Local Similarity 82.1%; Pred. No. 1,4e+03;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CCTCGCTAGTGTGACGACTCTCTGCG 28
|||||
276 CCTCCGTAGTGTGAGAGCTCTCTGCG 303

Db

RESULT 7
BI327791/c 723 bp mRNA linear EST 30-JUL-2001
LOCUS 602979537F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5132624 5',
DEFINITION mRNA sequence.
ACCESSION BI327791
VERSION BI327791.1 GI:15012448
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 723)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabers-rc@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1AM1325 row: 1 column: 09
High quality sequence start: 69.
High quality sequence stop: 609.
Location/Qualifiers
1..723
/organism="Mus musculus"
/mol_type="mRNA"
/strategy="FvB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5132624"
/lab_host="DH10B (T1 phage-resistant)"
/clone_id="NCI_CGAP_L19"
/note="Organ: liver; Vector: pCMV-Sport6; Site_1: NotI;

FEATURES
source

Site_2: Sal1; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.9 Kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT

197 a 176 c 211 g 139 t

ORIGIN

Query Match 66.7%; Score 20; DB 12; Length 723;
Best Local Similarity 82.1%; Pred. No. 1.5e+03;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 CCCCTGCTAGTGTGACCTCTCTGAG 28

718 CCATCGAAGCTGTGACCTCTCTGAG 691

RESULT 8

BY463661

LOCUS

BY463661 RIKEN full-length enriched, 10 days neonate heart Mus

DEFINITION

musculus cDNA clone K630144J20 3', mRNA sequence.

ACCESSION

BY463661 GI:26765911

VERSION

EST

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

REFERENCE

1 (bases 1 to 450)

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oshino, N., Salto, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Brad, D., Brusic, V., Chochla, C., Corbani,
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustlich, S., Hirokawa, N., Jackson, I. J.,
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
P. A., Maglott, D. R., Maltas, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, M. J., Perte, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou,
M., Shimada, K., Sultana, R., Takenaka, J., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlested, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wyszynski, B. A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carinici, P., Hayatsu, N., Hirozane-Kitahara, T., Kono, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Atsawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, N., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE

Analysis of the mouse transcriptome based on functional annotation

JOURNAL

Nature 420, 563-573 (2002)

MEDLINE

22354683

PUBMED

12466851

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Shohiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carinici, P., Fukuda, S., Hirozane,
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A.,
Miyazaki, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sasaki, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, E., Shiraki, T., Tagami,
M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct

FEATURES

source

1..450

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="K630144J20"

/tissue_type="heart"

/dev_stage="10 days neonate"

/clone_11b="RIKEN full-length enriched, 10 days neonate heart"

Location/Qualifiers

BASE COUNT

102 a 118 c 124 g 104 t 2 others

ORIGIN

Query Match 66.0%; Score 19.8; DB 13; Length 450;
Best Local Similarity 87.5%; Pred. No. 1.6e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

7 GTAGTGTACGACCTCTCTGACGG 30
303 GGAGTGTCTTCACACTCTCTGACGG 326

RESULT 9

CD475739/c

LOCUS

CD475739/c

DEFINITION

nad03-15me2-e07 Nad03 Nuphar advena cDNA clone nad03-15me2-e07 5',

RNA sequence.

CD475739

VERSION

CD475739.1 GI:31397007

KEYWORDS

EST.

SOURCE

Nuphar advena

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; stem Magnoliophyta; Nymphaeaceae;

Nuphar.

1 (bases 1 to 161)

Depamphilis, C., Carlson, J., Ma, H., Frohlich, M., Tanksey, S.,
Leebens-Mack, J., Field, D., Arrington, J., Zahn, L., Kong, H.,
Druckemiller, M., Landherr, L., Hu, Y., Ilut, D., Wall, K., Plock,
S., Chioresu, S., Albert, V., Doyle, J., Miller, W., Oppenheimer, D.,
Soltis, D., Soltis, P. and Theissen, G.

Generation of ESTs from early flower buds of Nuphar advena

Unpublished

Contact: Claude Depamphilis or James Leebens-Mack

Mueller Laboratory

Penn State University

208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn

State University, University Park, PA 16802, USA

Tel: 814 865 6413

Fax: 814 865 9131

Email: cw33@psu.edu or jhl10@psu.edu

The sequence provided is trimmed of vector and low quality regions.

Full sequence and original trace file are available from the Plant

Genome Network website (<http://pgrn.cornell.edu>)
 Place: nad03-15ms2 row: e column: 07
 Seq primer: M13F.

FEATURES

source

1.161
 /organism="Nuphar advena"
 /mol_type="mRNA"
 /db_xref="taxon:77108"
 /clone="nad03-15ms2-e07"
 /cissue_type="flower buds"
 /dev_stage="2.5mm"
 /lab_host="SOLR"
 /clone_lib="Nad03"
 /note="Vector: pBluescript SK (+/-); site 1: EcoRI;
 site 2: XhoI; Only floral buds with diameter of 2.5 mm of
 less were used for RNA isolation. This is a directionally
 cloned, non-normalized library. Avg. insert length: 1134;
 Primers: M13F and M13R; Antibiotic: 50 ug/ml Ampicillin;
 Primary titer: 286 pfu total; Amplified titer: 3.2E10
 pfu/ml; Mass Excised Titer: 5E10 total; This library has
 been generated by the Floral Genome Project (FGP). The
 Floral Genome Project is funded by NSF's Plant Genome
 Research Program (DBI-0115684). More information about the
 project can be obtained at <http://fgp.bio.psu.edu>"

BASE COUNT

40 a 47 c 45 g 29 t

ORIGIN

Query Match 65.3%; Score 19.6; DB 14; Length 161;
 Best Local Similarity 84.6%; Pred. No. 1.7e+03;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 5 GCCTAGTGTACGACCTCTGCAGG 30
 110 GCCTAGTGTACGACCTCTGCAGG 85

Db 110 GCCTAGTGTACGACCTCTGCAGG 85

RESULT 10 380 bp mRNA linear EST 27-NOV-2001
 BB872020
 LOCUS
 DEFINITION
 BB872020 RIKEN full-length enriched, 16 days neonate male
 diencephalon Mus musculus cDNA clone G630034A16 5', mRNA sequence.

ACCESSION
 BB872020
 VERSION
 BB872020.1 GI:17118230
 KEYWORDS
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus (house mouse)

REFERENCE
 AUTHORS
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 380)

Hayatsu, N., Arikawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
 Hayatsu, N., Hiraoka, T., Hirozane, T., Imotani, K., Ishii,
 Y., Ito, M., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Matsuyama, T.,
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T.,
 Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Seto, K.,
 Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa,
 A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T.,
 Watanabe, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.
 2001)

TITLE
 JOURNAL
 COMMENT

Unpublished
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
 M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)
 Wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Wachihi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura,
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
 Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 e mouse tissues.

FEATURES

source

1.380
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G630034A16"
 /sex="male"
 /cissue_type="dienecephalon"
 /dev_stage="16 days neonate"
 /clone_lib="RIKEN full-length enriched, 16 days neonate
 male dienecephalon"

BASE COUNT

51 a 97 c 129 g 103 t

ORIGIN

Query Match 64.7%; Score 19.4; DB 10; Length 380;
 Best Local Similarity 79.3%; Pred. No. 2.3e+03;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 1 CCTGCGTGTGTACGACCTCTGCAGG 29
 12 CCTGCGTGTGTACGACCTCTGCAGG 40

Db

RESULT 11 406 bp mRNA linear EST 24-DEC-1997
 AA709478
 LOCUS
 DEFINITION
 AA709478 v34a05.r1 Stratogene mouse heart (#937316) Mus musculus cDNA clone
 IMAGE:1224272 5', mRNA sequence.

ACCESSION
 AA709478
 VERSION
 AA709478.1 GI:2719336
 KEYWORDS
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus (house mouse)

REFERENCE
 AUTHORS
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 406)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisell, S., Kuwaba, T., Lacy, M., Le, M., Martin, J., Norris, M.,
 Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE
 JOURNAL
 COMMENT

Unpublished
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGS Consortium (info@image.llnl.gov) for further information.
 MGI:649864
 Seq primer: -28m3 rev1 ET from Amerham
 High quality sequence stop: 401.
 Location/Qualifiers

SOURCE

1. 406
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="IMAGE:1224272"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: heart; Vector: pBlueScript SK-; Site: 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGACGACG 3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'."

BASE COUNT 65 a 117 c 132 g 92 t

ORIGIN

Query Match 64.7%; Score 19.4; DB 9; Length 406;
Best Local Similarity 79.3%; Pred. No. 2.3e+03;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CCTGCTGAGTGTGACGACCTCTCCAG 29
||||| ||||| ||||| ||||| |||||
Db 168 CCTGTGAGTGTGACGACCTCTCCAG 196

RESULT 12
LOCUS BY680656 432 bp mRNA linear EST 16-DEC-2002
DEFINITION BY680656 RIKEN full-length enriched, 12.5 days embryo Rathke's
pouches Mus musculus cDNA clone K920033C15 3', mRNA sequence.
ACCESSION BY680656
VERSION BY680656.1 GI:27073409
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 432)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Saito, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,
Quackenbush, J., Schiml, L.M., Kampin, A., Matsuda, H., Batalov, S.,
Beisel, K.W., Blake, D.A., Bradt, D., Brusic, V., Chochia, C., Corbani,
L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,
A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Giesi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedierski, R.M.,
King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,
P.A., Maglott, D.R., Maltas, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontus, J.U., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring,
B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C.A., Secou,
M., Shimada, K., Sultana, R., Takenaki, Y., Taylor, M.S., Teasdale,
R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Waki, J., Aizawa, K.,
Atakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kigawa, I., Miyazaki, A., Sakai, K., Sakai, D., Shibata,
K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

TITLE
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 22354683
PUBMED 12466851

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome-gsc.riken.go.jp/
Aizawa, K., Akimura, T., Atakawa, T., Carninci, P., Fukuda, S., Hirozane,
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
M., Waki, K., Watanabe, A., Watanabe, M. and Hayashizaki, Y. Direct
Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse libraries.
Libraries were provided by Michelle Brinkmeier and Sally Camper (Dept. Human Genetics University of Michigan Medical School 4301
MSB 3 1500 W. Medical Center Dr. Ann Arbor, MI 48109-0638 USA)
whose assistance we gratefully acknowledge.
Please visit our web site (http://genome-gsc.riken.go.jp) for
further details.
FEATURES
source
1. 432
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD-1"
/db_xref="taxon:10090"
/clone="K920033C15"
/tissue_type="Rathke's pouches"
/dev_stage="12.5 days embryo"
/clone_lib="RIKEN full-length enriched, 12.5 days embryo
Rathke's pouches"
BASE COUNT 114 a 100 c 84 g 134 t

ORIGIN

Query Match 64.7%; Score 19.4; DB 13; Length 432;
Best Local Similarity 79.3%; Pred. No. 2.3e+03;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CCTGCTGAGTGTGACGACCTCTCCAG 29
||||| ||||| ||||| ||||| |||||
Db 118 CCTGCTGAGTGTGACGACCTCTCCAG 146

RESULT 13
LOCUS CB790399 438 bp mRNA linear EST 16-MAY-2003
DEFINITION AMGN:NC:00017-D1-A-W Rat pituitary (10472) Rattus norvegicus
cDNA clone nrip4-00017-d1 5', mRNA sequence.
ACCESSION CB790399
VERSION CB790399.1 GI:29878791
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 438)
 AUTHORS Amgen EST Program.
 TITLE Amgen Rat EST Program
 JOURNAL Unpublished
 COMMENT Contact: Dan Fitzpatrick
 Amgen, Inc
 One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
 Tel: 805 447-4861
 Plate: 00017 row: d column: 1.
 Location/Qualifiers
 1.438
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="nrp14-00017-d1"
 /clone_lib="W Rat pituitary (10472)"
 /note="Vector: pSPORT1; Site_1: SalI; Site_2: NcoI; W Rat pituitary adult female Wistar rat avg insert size 2.1 kb"

BASE COUNT 117 a 93 c 121 g 70 t
 ORIGIN

Query Match 64.7%; Score 19.4; DB 14; Length 438;
 Best Local Similarity 79.3%; Pred. No. 2.3e+03;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CCTGCGTAGTGTAGACACCTCTGCAGG 29
 |||||
 Db 331 CTCGTGAGTGTCTGGAGACCTCTGCAGG 359

RESULT 14
 LOCUS BB857961 474 bp mRNA linear EST 26-NOV-2001
 DEFINITION BB857961. RIKEN full-length enriched, B16 F10Y cells Mus musculus
 CDNA clone G370047P05 5', mRNA sequence.
 BB857961
 EST.
 BB857961.1 GI:17099415

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 474)
 Akimura,T., Arai,K., Carinci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hirao,Y., Hirozane,T., Imocani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Komori,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numata,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watanishi,A., Yasunishi,A., Yamamoto,M., and Hayashizaki,Y.
 RIKEN Encyclopedia of Mouse Full-length CDNA (Akimura,T., et al. 2001)

TITLE Unpublished
 JOURNAL Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Carinci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komori,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 waki,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,T., Matsuki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A., and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Komori,H., Fukunishi,Y., Shibata,K., Itoh,M., Carinci,P., Sugahara,Y., and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
 e mouse tissues.
 Location/Qualifiers
 1.474
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G370047P05"
 /clone_lib="B16 F10Y cells"
 /cell_type="B16 F10Y cells"
 /clone_lib="RIKEN full-length enriched, B16 F10Y cells"
 73 a 119 c 161 g 121 t

BASE COUNT
 ORIGIN

Query Match 64.7%; Score 19.4; DB 10; Length 474;
 Best Local Similarity 79.3%; Pred. No. 2.3e+03;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CCTGCGTAGTGTAGACACCTCTGCAGG 29
 |||||
 Db 37 CCTGTGAGTGTCTGGAGACCTCTGCAGG 65

RESULT 15
 LOCUS BF022130 531 bp mRNA linear EST 29-DEC-2000
 DEFINITION uy47b10.y1 NCI CGAP Lu30 Mus musculus cDNA clone IMAGE:3662683 5' similar to TR:Q9WTM1;Q9WTM1 MROCT., mRNA sequence.
 BF022130
 EST.
 BF022130.1 GI:10753462

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 531)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

TITLE Unpublished
 JOURNAL Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 image.llnl.gov/image/html/resources.shtml
 MG1:1423451
 Seq primer: -40RP from Gibco
 High quality sequence eLoc: 436.
 Location/Qualifiers
 1.531
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:3662683"
 /tissue_type="tumor, metastatic to mammary"
 /lab_host="DH10B"
 /clone_lib="NCI-CGAP Lu30"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NcoI; Site_2: SalI; transgenic model WNT-1, expression driven by

MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
GT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 66 a 184 c 169 g 112 t
ORIGIN

Query Match 64.7%; Score 19.4; DB 10; Length 531;
Best Local Similarity 79.3%; Pred. No. 2.4e+03;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Oy 1 CCTGCTAGTGTACGACCTCCTGCAGG 29
Db 409 CCTGTGAGTGGAGAGAGCTGCTGCAGG 437

Search completed: September 16, 2003, 20:39:57
Job time : 1522.42 secs

BEST AVAILABLE COPY

Wed Sep 17 10:20:10 2003

US-09-594-065-1.inpm

Schenker L.
09/15/03
Seq. IDs 1-4 w/notes
Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 18:28:15 ; Search time 1450.49 Seconds

(without alignments)
457.275 Million cell updates/sec

Title: US-09-594-065-1

Perfect score: 20

Sequence: 1 tcaccacgtcagcaccctc 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3363688 seqs, 1658189874 residues

Total number of hits satisfying chosen parameters: 66727376

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_Main:

1: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*

2: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*

3: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*

4: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*

5: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*

6: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*

7: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*

8: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*

9: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*

10: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*

11: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*

12: /cgn2_6/ptodata/2/pna/US092_COMB.seq.*

13: /cgn2_6/ptodata/2/pna/US093_COMB.seq.*

14: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*

15: /cgn2_6/ptodata/2/pna/US095_COMB.seq.*

16: /cgn2_6/ptodata/2/pna/US096_COMB.seq.*

17: /cgn2_6/ptodata/2/pna/US097_COMB.seq.*

18: /cgn2_6/ptodata/2/pna/US098_COMB.seq.*

19: /cgn2_6/ptodata/2/pna/US099_COMB.seq.*

20: /cgn2_6/ptodata/2/pna/US100_COMB.seq.*

21: /cgn2_6/ptodata/2/pna/US101_COMB.seq.*

22: /cgn2_6/ptodata/2/pna/US102_COMB.seq.*

23: /cgn2_6/ptodata/2/pna/US103_COMB.seq.*

24: /cgn2_6/ptodata/2/pna/US104_COMB.seq.*

25: /cgn2_6/ptodata/2/pna/US105_COMB.seq.*

26: /cgn2_6/ptodata/2/pna/US106_COMB.seq.*

27: /cgn2_6/ptodata/2/pna/US107_COMB.seq.*

28: /cgn2_6/ptodata/2/pna/US108_COMB.seq.*

29: /cgn2_6/ptodata/2/pna/US109_COMB.seq.*

30: /cgn2_6/ptodata/2/pna/US110_COMB.seq.*

31: /cgn2_6/ptodata/2/pna/US111_COMB.seq.*

32: /cgn2_6/ptodata/2/pna/US112_COMB.seq.*

33: /cgn2_6/ptodata/2/pna/US113_COMB.seq.*

34: /cgn2_6/ptodata/2/pna/US114_COMB.seq.*

35: /cgn2_6/ptodata/2/pna/US115_COMB.seq.*

36: /cgn2_6/ptodata/2/pna/US116_COMB.seq.*

37: /cgn2_6/ptodata/2/pna/US117_COMB.seq.*

38: /cgn2_6/ptodata/2/pna/US118_COMB.seq.*

39: /cgn2_6/ptodata/2/pna/US119_COMB.seq.*

40: /cgn2_6/ptodata/2/pna/US120_COMB.seq.*

41: /cgn2_6/ptodata/2/pna/US121_COMB.seq.*

42: /cgn2_6/ptodata/2/pna/US122_COMB.seq.*

43: /cgn2_6/ptodata/2/pna/US123_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	ID	Description
No.						
1	20	100.0	20	25	US-09-594-065-1	Sequence 1, Appli
2	20	100.0	74	25	US-09-594-065-6	Sequence 8, Appli
3	20	100.0	109	25	US-09-594-065-12	Sequence 12, Appli
4	20	100.0	109	25	US-09-594-065-13	Sequence 13, Appli

```

5      20      100.0      109      25      US-09-594-065-14      Sequence 14, Appl
6      20      100.0      481      34      US-09-821-837-9580      Sequence 9580, Ap
7      20      100.0      495      34      US-09-821-837-9581      Sequence 9581, Ap
8      20      100.0      2645      6      US-08-123-456-141      Sequence 141, App
9      20      100.0      2645      18      US-09-297-477A-141      Sequence 141, App
10     20      100.0      2645      43      US-09-994-404-141      Sequence 141, App
11     20      100.0      2713      13      US-08-804-439-13      Sequence 13, Appl
12     20      100.0      2713      19      US-09-301-390-13      Sequence 13, Appl
13     20      100.0      2713      19      US-09-338-328-13      Sequence 13, Appl
14     20      100.0      2715      96      US-60-413-956-11      Sequence 11, Appl
15     20      100.0      2841      6      US-08-123-456-78      Sequence 78, Appl
16     20      100.0      2841      18      US-09-297-477A-78      Sequence 78, Appl
17     20      100.0      2841      43      US-09-994-404-78      Sequence 78, Appl
18     20      100.0      117213      6      US-08-123-456-217      Sequence 217, App
19     20      100.0      117213      18      US-09-297-477A-217      Sequence 217, App
20     20      100.0      117213      43      US-09-994-404-217      Sequence 217, App
21     20      100.0      154746      1      PCT-US01-11372-8      Sequence 8, Appl1
22     20      100.0      154746      2      PCT-US01-11372-8      Sequence 8, Appl1
23     20      100.0      154746      34      US-09-827-688-8      Sequence 8, Appl1
24     20      100.0      34094      42      US-09-963-038A-1      Sequence 1, Appl1
25     18.4      92.0      34094      47      US-10-199-550-1      Sequence 1, Appl1
26     18.4      92.0      34094      48      US-10-245-603A-1      Sequence 1, Appl1
27     18.4      92.0      34094      62      US-60-081-882-1      Sequence 1, Appl1
28     17.4      87.0      302      83      US-60-288-292-11509      Sequence 11509, A
29     17.4      87.0      354      28      US-09-654-617-224251      Sequence 224251, A
30     17.4      87.0      354      30      US-09-684-016-224251      Sequence 224251, A
31     17.4      87.0      387      34      US-09-804-730-22459      Sequence 22459, A
32     17.4      87.0      387      72      US-60-189-657-22453      Sequence 22453, A
33     17.4      87.0      415      34      US-09-804-730-21290      Sequence 21290, A
34     17.4      87.0      415      72      US-60-189-657-21284      Sequence 21284, A
35     17.4      87.0      420      44      US-10-020-338-3102      Sequence 3102, Ap
36     17.4      87.0      428      34      US-09-804-730-20654      Sequence 20654, A
37     17.4      87.0      428      72      US-60-189-657-20648      Sequence 20648, A
38     17.4      87.0      471      71      US-60-172-376-1446      Sequence 1446, Ap
39     17.4      87.0      476      28      US-09-654-617-130864      Sequence 130864, A
40     17.4      87.0      476      30      US-09-684-016-130864      Sequence 130864, A
41     17.4      87.0      488      25      US-09-565-309A-1706      Sequence 1706, Ap
42     17.4      87.0      511      25      US-09-565-309A-44343      Sequence 44343, A
43     17.4      87.0      523      22      US-09-505-533-38163      Sequence 38163, A
44     17.4      87.0      523      34      US-09-819-091A-38163      Sequence 38163, A
45     17.4      87.0      541      71      US-60-172-519-8571      Sequence 8571, Ap

```

ALIGNMENTS

```

RESULT 1
; Sequence 1, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: Walkerpetch, Cindy
; APPLICANT: Dubois, Dwight
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 an
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Forward PCR primer
US-09-594-065-1

```

```

Query Match      100.0%; Score 20; DB 25; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 TCACCAACGCTGACGACCTTC 20
|||||

```

```

Db      1 TCACCAACGCTGACGACCTTC 20
RESULT 2
; Sequence 8, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: Walkerpetch, Cindy
; APPLICANT: Dubois, Dwight
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 a
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 74
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Forward primer for synthesis of IAC
US-09-594-065-8

```

```

Query Match      100.0%; Score 20; DB 25; Length 74;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 TCACCAACGCTGACGACCTTC 20
|||||
Db      10 TCACCAACGCTGACGACCTTC 29

```

```

RESULT 3
; Sequence 12, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: Walkerpetch, Cindy
; APPLICANT: Dubois, Dwight
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 a
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 109
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence amplified by PCR reaction for HSV type-1
US-09-594-065-12

```

```

Query Match      100.0%; Score 20; DB 25; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 TCACCAACGCTGACGACCTTC 20
|||||
Db      1 TCACCAACGCTGACGACCTTC 20

```

```

RESULT 4
; Sequence 13, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: Walkerpetch, Cindy
; APPLICANT: Dubois, Dwight
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 a
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065

```

;; CURRENT FILING DATE: 2000-06-14
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 13
;; LENGTH: 109
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Sequence amplified by PCR for HSV type-2
US-09-594-065-13

Query Match 100.0%; Score 20; DB 25; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCACCACCGTCGACACCTTC 20
|||
Db 1 TCACCACCGTCGACACCTTC 20

RESULT 5
US-09-594-065-14
;; Sequence 14, Application US/09594065
;; GENERAL INFORMATION:
;; APPLICANT: Walkerpeach, Cindy
;; APPLICANT: Dubois, Dwight
;; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 an
;; FILE REFERENCE: 25436/1280
;; CURRENT APPLICATION NUMBER: US/09/594,065
;; CURRENT FILING DATE: 2000-06-14
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 14
;; LENGTH: 109
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Sequence amplified by PCR for IAC
US-09-594-065-14

Query Match 100.0%; Score 20; DB 25; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCACCACCGTCGACACCTTC 20
|||
Db 1 TCACCACCGTCGACACCTTC 20

RESULT 6
US-09-821-837-9580/C
;; Sequence 9580, Application US/09821837
;; GENERAL INFORMATION:
;; APPLICANT: Gearing, David P.
;; APPLICANT: Holtzman, Douglas A.
;; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
;; FILE REFERENCE: 1600.2076-001
;; CURRENT APPLICATION NUMBER: US/09/821,837
;; CURRENT FILING DATE: 2001-03-29
;; PRIOR APPLICATION NUMBER: US 60/193,481
;; PRIOR FILING DATE: 2000-03-29
;; NUMBER OF SEQ ID NOS: 9928
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 9580
;; LENGTH: 481
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(481)
;; OTHER INFORMATION: n = A,T,C or G

US-09-821-837-9580

Query Match 100.0%; Score 20; DB 34; Length 481;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCACCACCGTCGACACCTTC 20
|||
Db 381 TCACCACCGTCGACACCTTC 362

RESULT 7
US-09-821-837-9581
;; Sequence 9581, Application US/09821837
;; GENERAL INFORMATION:
;; APPLICANT: Gearing, David P.
;; APPLICANT: Holtzman, Douglas A.
;; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
;; FILE REFERENCE: 1600.2076-001
;; CURRENT APPLICATION NUMBER: US/09/821,837
;; CURRENT FILING DATE: 2001-03-29
;; PRIOR APPLICATION NUMBER: US 60/193,481
;; PRIOR FILING DATE: 2000-03-29
;; NUMBER OF SEQ ID NOS: 9928
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 9581
;; LENGTH: 495
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(495)
;; OTHER INFORMATION: n = A,T,C or G
US-09-821-837-9581

Query Match 100.0%; Score 20; DB 34; Length 495;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCACCACCGTCGACACCTTC 20
|||
Db 159 TCACCACCGTCGACACCTTC 178

RESULT 8
US-08-123-456-141/C
;; Sequence 141, Application US/08123456
;; GENERAL INFORMATION:
;; APPLICANT: ESSER, KLAUS M.
;; APPLICANT: CHAN, JOHN Y.
;; APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
;; APPLICANT: DELVECCHIO, ALFRED MICHAEL
;; APPLICANT: DILLOM, SUSAN B.
;; APPLICANT: LEARY, JEFFREY JOSEPH
;; APPLICANT: SUTTON, DAVID
;; TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
;; FILE REFERENCE: P50583
;; CURRENT APPLICATION NUMBER: US/08/123,456
;; CURRENT FILING DATE: 1997-03-03
;; PRIOR APPLICATION NUMBER: US 60/030,279
;; PRIOR FILING DATE: 1999-11-04
;; PRIOR APPLICATION NUMBER: US 60/049,018
;; PRIOR FILING DATE: 1997-06-09
;; NUMBER OF SEQ ID NOS: 303
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 141
;; LENGTH: 2645
;; TYPE: DNA
;; ORGANISM: Herpes simplex
US-08-123-456-141

Query Match 100.0%; Score 20; DB 6; Length 2645;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCAACGCTGACGACCTTC 20
DB 814 TCACCAACGCTGACGACCTTC 795

RESULT 9
US-09-297-477A-141/c
Sequence 141, Application US/09297477A

GENERAL INFORMATION:
APPLICANT: ESSER, KLAUS M.
APPLICANT: CHAN, JOHN Y.
APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
APPLICANT: DELVECCHIO, ALFRED MICHAEL
APPLICANT: DILLON, SUSAN B.
APPLICANT: LEARY, JEFFREY JOSEPH
APPLICANT: SUTTON, DAVID
TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
FILE REFERENCE: P50583
CURRENT APPLICATION NUMBER: US/09/297,477A
CURRENT FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US 60/030,279
PRIOR FILING DATE: 1999-11-04
PRIOR APPLICATION NUMBER: US 60/049,018
PRIOR FILING DATE: 1997-06-09
NUMBER OF SEQ ID NOS: 303
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 141
LENGTH: 2645
TYPE: DNA
ORGANISM: Herpes simplex
US-09-297-477A-141

Query Match 100.0%; Score 20; DB 18; Length 2645;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCAACGCTGACGACCTTC 20
DB 814 TCACCAACGCTGACGACCTTC 795

RESULT 10
US-09-994-404-141/c
Sequence 141, Application US/09994404
GENERAL INFORMATION:
APPLICANT: ESSER, KLAUS M.
APPLICANT: CHAN, JOHN Y.
APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
APPLICANT: DELVECCHIO, ALFRED MICHAEL
APPLICANT: DILLON, SUSAN B.
APPLICANT: LEARY, JEFFREY JOSEPH
APPLICANT: SUTTON, DAVID
TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
FILE REFERENCE: P50583
CURRENT APPLICATION NUMBER: US/09/994,404
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/297,477
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US 60/030,279
PRIOR FILING DATE: 1999-11-04
PRIOR APPLICATION NUMBER: US 60/049,018
PRIOR FILING DATE: 1997-06-09
NUMBER OF SEQ ID NOS: 303
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 141
LENGTH: 2645
TYPE: DNA

ORGANISM: Herpes simplex
US-09-994-404-141

Query Match 100.0%; Score 20; DB 43; Length 2645;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCAACGCTGACGACCTTC 20
DB 814 TCACCAACGCTGACGACCTTC 795

RESULT 11
US-08-804-439-13
Sequence 13, Application US/08804439

GENERAL INFORMATION:
APPLICANT: ROSE, TIMOTHY M.
APPLICANT: BOSE, MARTIN L.
APPLICANT: STRAND, KURT
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSER: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,439
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: SCHIFF, J. MICHAEL
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.20
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2713 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-804-439-13

Query Match 100.0%; Score 20; DB 13; Length 2713;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCAACGCTGACGACCTTC 20
DB 1988 TCACCAACGCTGACGACCTTC 2007

RESULT 12
US-09-301-390-13
Sequence 13, Application US/09301390
GENERAL INFORMATION:
APPLICANT: ROSE, TIMOTHY M.
APPLICANT: BOSE, MARTIN L.
APPLICANT: STRAND, KURT
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 100

;; CORRESPONDENCE ADDRESS:
;; ADDRESSER: Morrison & Foerster
;; STREET: 755 Page Mill Road
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/301,390
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/720,229
;; FILING DATE: 26-SEP-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Schiff, J. Michael
;; REGISTRATION NUMBER: 40,253
;; REFERENCE/DOCKET NUMBER: 29938-20002.00
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 813-5600
;; TELEFAX: (415) 494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2713 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; US-09-301-390-13

Query Match 100.0%; Score 20; DB 19; Length 2713;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCACCGTCGACGCTTC 20
|||||
DB 1988 TCACCACCGTCGACGCTTC 2007

RESULT 13
US-09-338-326-13
;; Sequence 13, Application US/09338326
;; GENERAL INFORMATION:
;; APPLICANT: Rose, Timothy M.
;; APPLICANT: Bosch, Marix L.
;; APPLICANT: Strand, Kurt
;; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RRV/KSHV
;; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
;; NUMBER OF SEQUENCES: 100
;; CORRESPONDENCE ADDRESS:
;; ADDRESSER: Morrison & Foerster
;; STREET: 755 Page Mill Road
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/338,326
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/720,229
;; FILING DATE: 26-SEP-1996

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Schiff, J. Michael
;; REGISTRATION NUMBER: 40,253
;; REFERENCE/DOCKET NUMBER: 29938-20002.00
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 813-5600
;; TELEFAX: (415) 494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2713 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; US-09-338-326-13

Query Match 100.0%; Score 20; DB 19; Length 2713;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCACCGTCGACGCTTC 20
|||||
DB 1988 TCACCACCGTCGACGCTTC 2007

RESULT 14
US-60-412-956-11
;; Sequence 11, Application US/60412956
;; GENERAL INFORMATION:
;; APPLICANT: STYKES, KATHRYN F.
;; APPLICANT: STEWKE-HALE, KATHERINE
;; APPLICANT: JOHNSTON, STEPHEN ALBERT
;; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
;; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF THE
;; FILE REFERENCE: MCRO:002USP1
;; CURRENT APPLICATION NUMBER: US/60/412,956
;; CURRENT FILING DATE: 2002-05-23
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 11
;; LENGTH: 2715
;; TYPE: DNA
;; ORGANISM: HERPES VIRUS, TYPE 1
;; US-60-412-956-11

Query Match 100.0%; Score 20; DB 96; Length 2715;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCACCGTCGACGCTTC 20
|||||
DB 1991 TCACCACCGTCGACGCTTC 2010

RESULT 15
US-08-123-456-78
;; Sequence 78, Application US/08123456
;; GENERAL INFORMATION:
;; APPLICANT: ESSER, KLAUS M.
;; APPLICANT: CHAN, JOHN Y.
;; APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
;; APPLICANT: DELVECCHIO, ALFRED MICHAEL
;; APPLICANT: DILLON, SUSAN B.
;; APPLICANT: LEARY, JEFFREY JOSEPH
;; APPLICANT: SUTTON, DAVID
;; TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
;; TITLE OF INVENTION: SIMPLEX VIRUS TYPE-2
;; FILE REFERENCE: P50583
;; CURRENT APPLICATION NUMBER: US/08/123,456
;; CURRENT FILING DATE: 1997-03-03
;; PRIOR APPLICATION NUMBER: US 60/030,279
;; PRIOR FILING DATE: 1999-11-04

/ PRIOR APPLICATION NUMBER: US 60/049,018
/ PRIOR FILING DATE: 1997-06-09
/ NUMBER OF SEQ ID NOS: 303
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 78
/ LENGTH: 2841
/ TYPE: DNA
/ ORGANISM: Herpes simplex
/ FEATURE:
/ NAME/KEY: unknown
/ LOCATION: (2591)(2825)(2833)
/ OTHER INFORMATION:
US-08-123-456-78

Query Match 100.0%; Score 20; DB 6; Length 2841;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCACCAACCGTCACACCTTC 20
Db 2012 TCACCAACCGTCACACCTTC 2031

Search completed: September 16, 2003, 22:44:43
Job time : 1452.49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 16, 2003, 18:37:15 ; Search time 31.4851 Seconds
(without alignments)
535.803 Million cell updates/sec

Title: US-09-594-065-1

Perfect score: 20

Sequence: 1 tcaccacgcgtcagcacttc 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 606776 seqs, 421745864 residues

Total number of hits satisfying chosen parameters: 1213552

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_New.*

1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB_seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB_seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB_seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB_seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB_seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB_seq.*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB_seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.8	84.0	527	US-10-425-114A-27295	Sequence 27295, A
2	16.8	84.0	2086	US-10-425-114A-21287	Sequence 21287, A
3	16.8	84.0	7838	US-09-654-936A-147	Sequence 1467, Ap
4	16.8	84.0	91672	US-60-487-610-19444	Sequence 19444, A
5	16.8	84.0	2731748	US-10-297-465B-1	Sequence 1, Appli
6	16.4	82.0	25086	US-60-487-610-18609	Sequence 19609, A
7	16.4	82.0	25474	US-60-495-114-16832	Sequence 16832, A
8	16.4	82.0	207149	US-60-487-610-19377	Sequence 19377, A
9	16	80.0	614	US-10-425-114A-19377	Sequence 19377, A
10	15.8	79.0	179	PCT-US03-18714-26075	Sequence 26075, A
11	15.8	79.0	630	PCT-US02-37235-115	Sequence 115, App
12	15.8	79.0	749	US-10-425-114A-7593	Sequence 7593, Ap
13	15.8	79.0	772	US-10-425-114A-14861	Sequence 14861, A
14	15.8	79.0	784	US-10-425-114A-26570	Sequence 26570, A
15	15.8	79.0	919	US-10-425-114A-24003	Sequence 24003, A
16	15.8	79.0	2264	US-10-425-114A-19130	Sequence 19130, A
17	15.8	79.0	2533	US-10-425-114A-35600	Sequence 35600, A
18	15.8	79.0	2649	US-10-425-114A-24312	Sequence 24312, A
19	15.8	79.0	4000	US-09-976-858-281	Sequence 281, App
20	15.8	79.0	6014	US-60-490-890-2216	Sequence 2216, Ap
21	15.8	79.0	11740	US-60-490-890-706	Sequence 706, App
22	15.8	79.0	142005	US-60-487-610-19647	Sequence 19647, A
23	15.8	79.0	200820	US-60-495-114-16936	Sequence 16936, A
24	15.8	79.0	4813087	US-09-947-914-75	Sequence 75, Appl
25	15.4	77.0	1271	US-10-425-114A-10829	Sequence 10829, A
26	15.4	77.0	1371	US-10-425-114A-11773	Sequence 11773, A

27	15.4	77.0	1972	6	US-10-425-114A-20032	Sequence 20032, A
28	15.4	77.0	2023	6	US-10-425-114A-10611	Sequence 10611, A
29	15.4	77.0	76798	7	US-60-490-890-1467	Sequence 1467, Ap
30	15.4	77.0	201007	7	US-60-487-610-19435	Sequence 19435, A
31	15.4	77.0	255673	7	US-60-495-114-15366	Sequence 16366, A
32	15.4	77.0	255673	7	US-60-495-114-15366	Sequence 16366, A
33	15.4	77.0	513030	7	US-60-487-610-19737	Sequence 19737, A
34	15.4	77.0	2731748	6	US-10-297-465B-1	Sequence 1, Appli
35	15.2	76.0	201	7	US-60-487-610-40113	Sequence 40113, A
36	15.2	76.0	201	7	US-60-487-610-40113	Sequence 40113, A
37	15.2	76.0	201	7	US-60-487-610-40115	Sequence 40115, A
38	15.2	76.0	201	7	US-60-487-610-40116	Sequence 40116, A
39	15.2	76.0	201	7	US-60-487-610-58562	Sequence 58562, A
40	15.2	76.0	201	7	US-60-487-610-69696	Sequence 69696, A
41	15.2	76.0	201	7	US-60-487-610-69700	Sequence 69700, A
42	15.2	76.0	201	7	US-60-487-610-69730	Sequence 69730, A
43	15.2	76.0	201	7	US-60-487-610-69747	Sequence 69747, A
44	15.2	76.0	201	7	US-60-487-610-69753	Sequence 69753, A
45	15.2	76.0	201	7	US-60-487-610-70283	Sequence 70283, A

ALIGNMENTS

```
RESULT 1
US-10-425-114A-27295/c
; Sequence 27295, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 27295
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURES:
; OTHER INFORMATION: Clone ID: LIB4729-056-B3_FLI
US-10-425-114A-27295

Query Match      84.0% Score 16.8; DB 6; Length 527;
Best Local Similarity 90.0% Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 TCACCACCGTCAGCAGCTTC 20
        |||||
DB      264 TCACCACCGTCAGCAGCTTC 245

RESULT 2
US-10-425-114A-21287/c
; Sequence 21287, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
```

SEQ ID NO 21287
LENGTH: 2086
TYPE: DNA
ORGANISM: Zee mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3279-218-A5_FLI
US-10-425-114A-21287

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 6; Length 2086;
90.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCACCACTGTCAGACCTTC 20
Db 620 TCACCACTGTCAGACCTTC 601

RESULT 3
US-09-654-936A-147/c
Sequence 147, Application US/09654936A
GENERAL INFORMATION:
APPLICANT: Tang, X. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhou, Ping
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Zhao, Qing A.
APPLICANT: Xue, Aidong J.
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 787CIP2C
CURRENT APPLICATION NUMBER: US/09/654,936A
CURRENT FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 164
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 147
LENGTH: 7838
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (673)..(2616)
US-09-654-936A-147

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 5; Length 7838;
90.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCACCACTGTCAGACCTTC 20
Db 6499 TCACCACTGTCAGACCTTC 6480

RESULT 4
US-60-487-610-19444
Sequence 19444, Application US/60487610
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: HUANG, Hongjin
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001469
CURRENT APPLICATION NUMBER: US/60/487,610
CURRENT FILING DATE: 2003-07-17

NUMBER OF SEQ ID NOS: 97101
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19444
LENGTH: 91672
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(91672)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-60-487-610-19444

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 7; Length 91672;
90.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCACCACTGTCAGACCTTC 20
Db 19866 TCACCACTGTCAGACCTTC 19885

RESULT 5
US-10-297-465B-1/c
Sequence 1, Application US/10297465B
GENERAL INFORMATION:
APPLICANT: Simpson, Andrew
APPLICANT: Reinach, Fernando
APPLICANT: Setubal, Joao
APPLICANT: Mediane, Joao
APPLICANT: Arruda, Paulo
TITLE OF INVENTION: Isolated Genome of *Xylella fastidiosa* and Uses Thereof
FILE REFERENCE: FAPESP 202 US (10213376)
CURRENT APPLICATION NUMBER: US/10/297,465B
CURRENT FILING DATE: 2003-08-25
PRIOR APPLICATION NUMBER: PCT/IB01/01618
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: 60/209,906
PRIOR FILING DATE: 2001-06-17
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 2731748
TYPE: DNA
ORGANISM: *Xylella fastidiosa*
US-10-297-465B-1

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 6; Length 2731748;
90.0%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCACCACTGTCAGACCTTC 20
Db 723037 TCACCACTGTCAGACCTTC 723018

RESULT 6
US-60-487-610-19609
Sequence 19609, Application US/60487610
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: HUANG, Hongjin
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001469
CURRENT APPLICATION NUMBER: US/60/487,610
CURRENT FILING DATE: 2003-07-17
NUMBER OF SEQ ID NOS: 97101
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19609
LENGTH: 25086
TYPE: DNA
ORGANISM: Homo sapiens

US-60-487-610-19609

Query Match 82.0%; Score 16.4; DB 7; Length 25086;
Best Local Similarity 94.4%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CACCACCGTCGACCTT 19
|||||
DB 203 CACCACCATCAGACCTT 220

RESULT 7

US-60-495-114-16832
; Sequence 16832, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO01480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16832
; LENGTH: 25474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(25474)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-60-495-114-16832

Query Match 82.0%; Score 16.4; DB 7; Length 25474;
Best Local Similarity 94.4%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CACCACCGTCGACCTT 19
|||||
DB 298 CACCACCATCAGACCTT 315

RESULT 8

US-60-487-610-19377/C
; Sequence 19377, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INJECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19377
; LENGTH: 207149
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(207149)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-60-487-610-19377

Query Match 82.0%; Score 16.4; DB 7; Length 207149;
Best Local Similarity 94.4%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCACCGTCGACCTT 18
|||||

DB 86537 TCACCACCATCAGACCTT 86520

RESULT 9

US-10-425-114A-19377/C
; Sequence 19377, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabeska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 19377
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3146-010-H6_FLI
US-10-425-114A-19377

Query Match 80.0%; Score 16; DB 6; Length 614;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CACCACCGTCGACCTT 17
|||||
DB 213 CACCACCGTCGACCTT 198

RESULT 10

PCT-US03-18714-26075/C
; Sequence 26075, Application PC/TUS0318714
; GENERAL INFORMATION:
; APPLICANT: Reguime Corporation
; TITLE OF INVENTION: Functional Sites
; FILE REFERENCE: 11207-029-228
; CURRENT APPLICATION NUMBER: PCT/US03/18714
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/387,887
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 60/387,910
; PRIOR FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 51999
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26075
; LENGTH: 179
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US03-18714-26075

Query Match 79.0%; Score 15.8; DB 1; Length 179;
Best Local Similarity 89.5%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCACCGTCGACCTT 20
|||||
DB 29 CTCACCCGTCGACCTT 11

RESULT 11

PCT-US02-37235-115
; Sequence 115, Application PC/TUS0237235
; GENERAL INFORMATION:
; APPLICANT: Handfield, Martin
; APPLICANT: Hillman, Jeffrey
; APPLICANT: Proguiske-Fox, Ann

```
/ TITLE OF INVENTION: Identification of Actinobacillus actinomycetemcomitans Antigens #
/ FILE REFERENCE: MBH01-662A
/ CURRENT APPLICATION NUMBER: PCT/US02/37235
/ NUMBER OF SEQ ID NOS: 234
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 115
/ LENGTH: 630
/ TYPE: DNA
/ ORGANISM: Actinobacillus actinomycetemcomitans
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(630)
PCT-US02-37235-115

Query Match          79.0%; Score 15.8; DB 1; Length 630;
Best Local Similarity 89.5%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 TCACGACCGTCAGCACCCTT 19
DB      386 TCGCCACCGTTACGACCTT 404

RESULT 12
/ Sequence 7593, Application US/10425114A
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114A
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 7593
/ LENGTH: 749
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: 700674452_FLI
US-10-425-114A-7593

Query Match          79.0%; Score 15.8; DB 6; Length 749;
Best Local Similarity 89.5%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 CACGACCGTCAGCACCCTT 20
DB      584 CACCACCATCACCACCTTC 566

RESULT 13
/ Sequence 14861, Application US/10425114A
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114A
/ CURRENT FILING DATE: 2003-04-28
```

```
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 14861
/ LENGTH: 772
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3029-011-F8_FLI
US-10-425-114A-14861

Query Match          79.0%; Score 15.8; DB 6; Length 772;
Best Local Similarity 89.5%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 CACGACCGTCAGCACCCTT 20
DB      608 CACCACCATCACCACCTTC 590

RESULT 14
/ Sequence 26570, Application US/10425114A
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114A
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 26570
/ LENGTH: 784
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB4315-040-D4_FLI
US-10-425-114A-26570

Query Match          79.0%; Score 15.8; DB 6; Length 784;
Best Local Similarity 89.5%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 CACGACCGTCAGCACCCTT 20
DB      147 CACCACCTTCAGCGCCTTC 129

RESULT 15
/ Sequence 24003, Application US/10425114A
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114A
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 24003
/ LENGTH: 919
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3606-055-G3_FLI
```

US-10-425-114A-24003

Query Match 79.0%; Score 15.8; DB 6; Length 919;

Best Local Similarity 89.5%; Pred. No. 3e+02; Mismatches 2; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TCACGACCGTCAGACCTT 19

Db 634 TCACCTCCTTCAGACCTT 616

Search completed: September 16, 2003, 22:47:51
Job time : 41.4851 secs

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 16, 2003, 18:28:15 ; Search time 1305.45 Seconds

(Without alignments)
457.275 Million cell updates/sec

Title: US-09-594-065-2

Perfect score: 18

Sequence: 1 agcagggccgctgccttg 18

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3363688 seqs, 1658189874 residues

Total number of hits satisfying chosen parameters: 66727376

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/2/pna/PCRTUS_COMB.seq.*

2: /cgn2_6/ptodata/2/pna/PCRTUS_COMB.seq.*

3: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*

4: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*

5: /cgn2_6/ptodata/2/pna/US08_COMB.seq.*

6: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*

7: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*

8: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*

9: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*

10: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*

11: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*

12: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*

13: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*

14: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*

15: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*

16: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*

17: /cgn2_6/ptodata/2/pna/US092A_COMB.seq.*

18: /cgn2_6/ptodata/2/pna/US092B_COMB.seq.*

19: /cgn2_6/ptodata/2/pna/US093A_COMB.seq.*

20: /cgn2_6/ptodata/2/pna/US093B_COMB.seq.*

21: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*

22: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*

23: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*

24: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*

25: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*

26: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*

27: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*

28: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*

29: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*

30: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*

31: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*

32: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*

33: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*

34: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*

35: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*

36: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*

37: /cgn2_6/ptodata/2/pna/US098D_COMB.seq.*

38: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*

39: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*

40: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*

41: /cgn2_6/ptodata/2/pna/US099D_COMB.seq.*

42: /cgn2_6/ptodata/2/pna/US099E_COMB.seq.*

43: /cgn2_6/ptodata/2/pna/US099F_COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US100A_COMB.seq.*

45: /cgn2_6/ptodata/2/pna/US100B_COMB.seq.*

46: /cgn2_6/ptodata/2/pna/US101A_COMB.seq.*

47: /cgn2_6/ptodata/2/pna/US101B_COMB.seq.*

48: /cgn2_6/ptodata/2/pna/US102A_COMB.seq.*

49: /cgn2_6/ptodata/2/pna/US102B_COMB.seq.*

50: /cgn2_6/ptodata/2/pna/US103A_COMB.seq.*

51: /cgn2_6/ptodata/2/pna/US103B_COMB.seq.*

52: /cgn2_6/ptodata/2/pna/US104A_COMB.seq.*

53: /cgn2_6/ptodata/2/pna/US104B_COMB.seq.*

54: /cgn2_6/ptodata/2/pna/US600A_COMB.seq.*

55: /cgn2_6/ptodata/2/pna/US600B_COMB.seq.*

56: /cgn2_6/ptodata/2/pna/US600C_COMB.seq.*

57: /cgn2_6/ptodata/2/pna/US600D_COMB.seq.*

58: /cgn2_6/ptodata/2/pna/US600E_COMB.seq.*

59: /cgn2_6/ptodata/2/pna/US600F_COMB.seq.*

60: /cgn2_6/ptodata/2/pna/US600G_COMB.seq.*

61: /cgn2_6/ptodata/2/pna/US600H_COMB.seq.*

62: /cgn2_6/ptodata/2/pna/US600I_COMB.seq.*

63: /cgn2_6/ptodata/2/pna/US600J_COMB.seq.*

64: /cgn2_6/ptodata/2/pna/US600K_COMB.seq.*

65: /cgn2_6/ptodata/2/pna/US600L_COMB.seq.*

66: /cgn2_6/ptodata/2/pna/US600M_COMB.seq.*

67: /cgn2_6/ptodata/2/pna/US600N_COMB.seq.*

68: /cgn2_6/ptodata/2/pna/US600O_COMB.seq.*

69: /cgn2_6/ptodata/2/pna/US600P_COMB.seq.*

70: /cgn2_6/ptodata/2/pna/US600Q_COMB.seq.*

71: /cgn2_6/ptodata/2/pna/US600R_COMB.seq.*

72: /cgn2_6/ptodata/2/pna/US600S_COMB.seq.*

73: /cgn2_6/ptodata/2/pna/US600T_COMB.seq.*

74: /cgn2_6/ptodata/2/pna/US600U_COMB.seq.*

75: /cgn2_6/ptodata/2/pna/US600V_COMB.seq.*

76: /cgn2_6/ptodata/2/pna/US600W_COMB.seq.*

77: /cgn2_6/ptodata/2/pna/US600X_COMB.seq.*

78: /cgn2_6/ptodata/2/pna/US600Y_COMB.seq.*

79: /cgn2_6/ptodata/2/pna/US600Z_COMB.seq.*

80: /cgn2_6/ptodata/2/pna/US600A_COMB.seq.*

81: /cgn2_6/ptodata/2/pna/US600B_COMB.seq.*

82: /cgn2_6/ptodata/2/pna/US600C_COMB.seq.*

83: /cgn2_6/ptodata/2/pna/US600D_COMB.seq.*

84: /cgn2_6/ptodata/2/pna/US600E_COMB.seq.*

85: /cgn2_6/ptodata/2/pna/US600F_COMB.seq.*

86: /cgn2_6/ptodata/2/pna/US600G_COMB.seq.*

87: /cgn2_6/ptodata/2/pna/US600H_COMB.seq.*

88: /cgn2_6/ptodata/2/pna/US600I_COMB.seq.*

89: /cgn2_6/ptodata/2/pna/US600J_COMB.seq.*

90: /cgn2_6/ptodata/2/pna/US600K_COMB.seq.*

91: /cgn2_6/ptodata/2/pna/US600L_COMB.seq.*

92: /cgn2_6/ptodata/2/pna/US600M_COMB.seq.*

93: /cgn2_6/ptodata/2/pna/US600N_COMB.seq.*

94: /cgn2_6/ptodata/2/pna/US600O_COMB.seq.*

95: /cgn2_6/ptodata/2/pna/US600P_COMB.seq.*

96: /cgn2_6/ptodata/2/pna/US600Q_COMB.seq.*

97: /cgn2_6/ptodata/2/pna/US600R_COMB.seq.*

98: /cgn2_6/ptodata/2/pna/US600S_COMB.seq.*

99: /cgn2_6/ptodata/2/pna/US600T_COMB.seq.*

100: /cgn2_6/ptodata/2/pna/US600U_COMB.seq.*

101: /cgn2_6/ptodata/2/pna/US600V_COMB.seq.*

102: /cgn2_6/ptodata/2/pna/US600W_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	18	100.0	18	25	US-09-594-065-2	Sequence 2, Appl
2	18	100.0	73	25	US-09-594-065-9	Sequence 9, Appl
3	18	100.0	109	25	US-09-594-065-12	Sequence 12, Appl
4	18	100.0	109	25	US-09-594-065-13	Sequence 13, Appl

```

c 5 18 100.0 109 25 US-09-594-065-14 Sequence 14, Appl
6 18 100.0 481 34 US-09-821-837-9580 Sequence 9580, Ap
c 7 18 100.0 495 34 US-09-821-837-9581 Sequence 9581, Ap
8 18 100.0 2645 6 US-08-123-456-141 Sequence 141, App
9 18 100.0 2645 18 US-09-297-477A-141 Sequence 141, App
10 18 100.0 2645 43 US-09-994-404-141 Sequence 141, App
c 11 18 100.0 2713 13 US-08-804-439-13 Sequence 13, Appl
c 12 18 100.0 2713 13 US-09-301-390-13 Sequence 13, Appl
c 13 18 100.0 2713 19 US-09-338-326-13 Sequence 13, Appl
c 14 18 100.0 2715 96 US-60-412-956-11 Sequence 11, Appl
c 15 18 100.0 2841 6 US-08-123-456-78 Sequence 78, Appl
c 16 18 100.0 2841 18 US-09-297-477A-78 Sequence 78, Appl
c 17 18 100.0 2841 43 US-09-994-404-78 Sequence 78, Appl
c 18 100.0 2943 10 US-08-541-878-7 Sequence 7, Appl
19 18 100.0 117213 6 US-08-123-456-217 Sequence 217, App
20 18 100.0 117213 18 US-09-297-477A-217 Sequence 217, App
21 18 100.0 117213 43 US-09-994-404-217 Sequence 217, App
22 18 100.0 154746 1 PCT-US01-11372-8 Sequence 8, Appl1
23 18 100.0 154746 2 PCT-US01-11372-8 Sequence 8, Appl1
24 18 100.0 154746 34 US-09-827-688-8 Sequence 1201, Ap
c 25 17 94.4 855 51 US-10-366-683-1201 Sequence 1201, Ap
26 17 94.4 855 52 US-10-419-128-1201 Sequence 1201, Ap
c 27 17 94.4 1431 1 PCT-US02-03987-7812 Sequence 7812, Ap
c 28 17 94.4 1431 34 US-09-815-242-7812 Sequence 7812, Ap
c 29 17 94.4 1431 45 US-10-072-851-7812 Sequence 7812, Ap
c 30 17 94.4 1431 49 US-10-282-122A-30286 Sequence 30286, A
c 31 17 94.4 1491 51 US-10-366-683-1243 Sequence 1243, Ap
c 32 17 94.4 1551 52 US-10-419-128-1243 Sequence 1243, Ap
c 33 17 94.4 1551 51 US-10-366-683-1262 Sequence 1362, Ap
c 34 17 94.4 1551 52 US-10-419-128-1162 Sequence 1362, Ap
c 35 17 94.4 6244 9 US-08-479-306-8 Sequence 8, Appl1
c 36 17 94.4 6244 18 US-09-281-674-8 Sequence 8, Appl1
c 37 17 94.4 6244 33 US-09-777-317-8 Sequence 8, Appl1
c 38 17 94.4 6244 37 US-09-892-227-8 Sequence 8, Appl1
c 39 17 94.4 6244 62 US-60-082-302-601 Sequence 601, App
c 40 17 94.4 10351 62 US-60-082-302-601 Sequence 601, App
c 41 16.4 91.1 19 1 PCT-US02-25943-35472 Sequence 35472, A
c 42 16.4 91.1 19 48 US-10-227-565-35472 Sequence 35472, A
c 43 16.4 91.1 19 51 US-10-367-832A-35472 Sequence 35472, A
c 44 16.4 91.1 201 100 US-60-452-680-71403 Sequence 71403, A
c 45 16.4 91.1 201 100 US-60-452-680-71404 Sequence 71404, A

```

ALIGNMENTS

```

RESULT 1
US-09-594-065-2
; Sequence 2, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: WalkerPeach, Cindy
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 an
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Reverse PCR primer
US-09-594-065-2

```

```

Query Match 100.0%; Score 18; DB 25; Length 18;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 AGCAGCCGCTGTCTTG 18
|||||

```

```

Db 1 AGCAGCCGCTGTCTTG 18
RESULT 2
US-09-594-065-9
; Sequence 9, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: WalkerPeach, Cindy
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 a
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 73
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Reverse primer for synthesis of IAC
US-09-594-065-9

```

```

Query Match 100.0%; Score 18; DB 25; Length 73;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 AGCAGCCGCTGTCTTG 18
|||||
10 AGCAGCCGCTGTCTTG 27

```

```

RESULT 3
US-09-594-065-12/C
; Sequence 12, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: WalkerPeach, Cindy
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 an
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 109
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence amplified by PCR reaction for HSV type-1
US-09-594-065-12

```

```

Query Match 100.0%; Score 18; DB 25; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 109 AGCAGCCGCTGTCTTG 92
|||||

```

```

RESULT 4
US-09-594-065-13/C
; Sequence 13, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: WalkerPeach, Cindy
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 an
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065

```



```

; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 109
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence amplified by PCR for HSV type-2
US-09-594-065-13

```

```

Query Match          100.0%; Score 18; DB 25; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AGCAGGCCGCTGTCTCTTG 18
        |||||||
Db      109 AGCAGGCCGCTGTCTCTTG 92

```

```

RESULT 5
US-09-594-065-14/c
; Sequence 14, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: WalkerPeach, Cindy
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 an
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 109
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence amplified by PCR for IAC
US-09-594-065-14

```

```

Query Match          100.0%; Score 18; DB 25; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AGCAGGCCGCTGTCTCTTG 18
        |||||||
Db      109 AGCAGGCCGCTGTCTCTTG 92

```

```

RESULT 6
US-09-821-837-9580
; Sequence 9580, Application US/09821837
; GENERAL INFORMATION:
; APPLICANT: Geating, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2076-001
; CURRENT APPLICATION NUMBER: US/09/821,837
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/193,481
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 9928
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9580
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(481)
; OTHER INFORMATION: n = A,T,C or G

```

```

US-09-821-837-9580

```

```

Query Match          100.0%; Score 18; DB 34; Length 481;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AGCAGGCCGCTGTCTCTTG 18
        |||||||
Db      273 AGCAGGCCGCTGTCTCTTG 290

```

```

RESULT 7
US-09-821-837-9581/c
; Sequence 9581, Application US/09821837
; GENERAL INFORMATION:
; APPLICANT: Geating, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2076-001
; CURRENT APPLICATION NUMBER: US/09/821,837
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/193,481
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 9928
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9581
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(495)
; OTHER INFORMATION: n = A,T,C or G
US-09-821-837-9581

```

```

Query Match          100.0%; Score 18; DB 34; Length 495;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AGCAGGCCGCTGTCTCTTG 18
        |||||||
Db      267 AGCAGGCCGCTGTCTCTTG 250

```

```

RESULT 8
US-08-123-456-141
; Sequence 141, Application US/08123456
; GENERAL INFORMATION:
; APPLICANT: ESSER, KLAUS M.
; APPLICANT: CHAN, JOHN Y.
; APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
; APPLICANT: DELVECCHIO, ALFRED MICHAEL
; APPLICANT: DILLON, SUSAN B.
; APPLICANT: LEARY, JEFFREY JOSEPH
; APPLICANT: SUTTON, DAVID
; TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
; FILE REFERENCE: P50583
; CURRENT APPLICATION NUMBER: US/08/123,456
; CURRENT FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: US 60/030,279
; PRIOR FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: US 60/049,018
; PRIOR FILING DATE: 1997-06-09
; NUMBER OF SEQ ID NOS: 303
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 141
; LENGTH: 2645
; TYPE: DNA
; ORGANISM: Herpes simplex
US-08-123-456-141

```

Query Match 100.0%; Score 18; DB 6; Length 2645;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGCAGCGCCGCTGTCTTG 18
|||||
DB 706 AGCAGCGCCGCTGTCTTG 723

RESULT 9

US-09-297-477A-141
Sequence 141, Application US/09297477A
GENERAL INFORMATION:
APPLICANT: ESSER, KLAUS M.
APPLICANT: CHAN, JOHN Y.
APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
APPLICANT: DILLOM, SUSAN B.
APPLICANT: DEVECCHIO, ALFRED MICHAEL
APPLICANT: LEARY, JEFFREY JOSEPH
APPLICANT: SUTTON, DAVID
TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
TITLE OF INVENTION: SIMPLEX VIRUS TYPE-2
FILE REFERENCE: P50583
CURRENT APPLICATION NUMBER: US/09/297,477A
CURRENT FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US 60/030,279
PRIOR FILING DATE: 1999-11-04
PRIOR APPLICATION NUMBER: US 60/049,018
PRIOR FILING DATE: 1997-06-09
NUMBER OF SEQ ID NOS: 303
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 141
LENGTH: 2645
TYPE: DNA
ORGANISM: Herpes simplex
US-09-297-477A-141

Query Match 100.0%; Score 18; DB 18; Length 2645;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGCAGCGCCGCTGTCTTG 18
|||||
DB 706 AGCAGCGCCGCTGTCTTG 723

RESULT 10

US-09-994-404-141
Sequence 141, Application US/09994404
GENERAL INFORMATION:
APPLICANT: ESSER, KLAUS M.
APPLICANT: CHAN, JOHN Y.
APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
APPLICANT: DEVECCHIO, ALFRED MICHAEL
APPLICANT: DILLOM, SUSAN B.
APPLICANT: LEARY, JEFFREY JOSEPH
APPLICANT: SUTTON, DAVID
TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
TITLE OF INVENTION: SIMPLEX VIRUS TYPE-2
FILE REFERENCE: P50583
CURRENT APPLICATION NUMBER: US/09/994,404
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/297,477
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US 60/030,279
PRIOR FILING DATE: 1999-11-04
PRIOR APPLICATION NUMBER: US 60/049,018
PRIOR FILING DATE: 1997-06-09
NUMBER OF SEQ ID NOS: 303
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 141
LENGTH: 2645
TYPE: DNA

ORGANISM: Herpes simplex
US-09-994-404-141

Query Match 100.0%; Score 18; DB 43; Length 2645;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGCAGCGCCGCTGTCTTG 18
|||||
DB 706 AGCAGCGCCGCTGTCTTG 723

RESULT 11

US-08-804-439-13/C
Sequence 13, Application US/08804439
GENERAL INFORMATION:
APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Marnix L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSER: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,439
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schief, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2713 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-804-439-13

Query Match 100.0%; Score 18; DB 13; Length 2713;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGCAGCGCCGCTGTCTTG 18
|||||
DB 2096 AGCAGCGCCGCTGTCTTG 2079

RESULT 12

US-09-301-390-13/C
Sequence 13, Application US/09301390
GENERAL INFORMATION:
APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Marnix L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 100

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Morrison & Foerster
;; STREET: 755 Page Mill Road
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/301,390
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/720,229
;; FILING DATE: 26-SEP-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Schiff, J. Michael
;; REGISTRATION NUMBER: 40,253
;; REFERENCE/DOCKET NUMBER: 29938-20002.00
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 813-5600
;; TELEFAX: (415) 494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2713 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; US-09-301-390-13

Query Match 100.0%; Score 18; DB 19; Length 2713;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGCTCTTG 18
DB 2096 AGCAGCGCGCTGCTCTTG 2079

RESULT 13
US-09-338-326-13/C
;; Sequence 13, Application US/09338326
;; GENERAL INFORMATION:
;; APPLICANT: Rose, Timothy M.
;; APPLICANT: Bosch, Martin L.
;; APPLICANT: Strand, Kurt
;; TITLE OF INVENTION: GLYCOPROTEIN B OF THE PHV/KSHV
;; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
;; NUMBER OF SEQUENCES: 100
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Morrison & Foerster
;; STREET: 755 Page Mill Road
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/338,326
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/720,229
;; FILING DATE: 26-SEP-1996

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Schiff, J. Michael
;; REGISTRATION NUMBER: 40,253
;; REFERENCE/DOCKET NUMBER: 29938-20002.00
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 813-5600
;; TELEFAX: (415) 494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2713 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; US-09-338-326-13

Query Match 100.0%; Score 18; DB 19; Length 2713;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGCTCTTG 18
DB 2096 AGCAGCGCGCTGCTCTTG 2079

RESULT 14
US-60-412-956-11/C
;; Sequence 11, Application US/60412956
;; GENERAL INFORMATION:
;; APPLICANT: SYKES, KATHRYN F.
;; APPLICANT: STEWKE-HALE, KATHERINE
;; APPLICANT: JOHNSTON, STEPHEN ALBERT
;; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
;; TITLE OF INVENTION: HERPESVIRUS FAMILY
;; FILE REFERENCE: MICRO:00205P1
;; CURRENT APPLICATION NUMBER: US/60/412,956
;; CURRENT FILING DATE: 2002-09-23
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 11
;; LENGTH: 2715
;; TYPE: DNA
;; ORGANISM: HERPES VIRUS, TYPE 1
;; US-60-412-956-11

Query Match 100.0%; Score 18; DB 96; Length 2715;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGCTCTTG 18
DB 2099 AGCAGCGCGCTGCTCTTG 2082

RESULT 15
US-08-123-456-78/C
;; Sequence 78, Application US/08123456
;; GENERAL INFORMATION:
;; APPLICANT: ESSER, KLAUS M.
;; APPLICANT: CHAN, JOHN Y.
;; APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
;; APPLICANT: DELVECCHIO, ALFRED MICHAEL
;; APPLICANT: DILLON, SUSAN B.
;; APPLICANT: LEAHY, JEFFREY JOSEPH
;; APPLICANT: SUTTON, DAVID
;; TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
;; TITLE OF INVENTION: SIMPLEX VIRUS TYPE-2
;; FILE REFERENCE: P50583
;; CURRENT APPLICATION NUMBER: US/08/123,456
;; CURRENT FILING DATE: 1997-03-03
;; PRIOR APPLICATION NUMBER: US 60/030,279
;; PRIOR FILING DATE: 1999-11-04

PRIOR APPLICATION NUMBER: US 60/049,018
PRIOR FILING DATE: 1997-06-09
NUMBER OF SEQ ID NOS: 303
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 78
LENGTH: 2841
TYPE: DNA
ORGANISM: Herpes simplex
FEATURE:
NAME/KEY: unknown
LOCATION: (2591)(2825)(2833)
OTHER INFORMATION:
US-08-123-456-78

Query Match 100.0%; Score 18; DB 6; Length 2841;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCAGGCCGCTGTCTTG 18
|||||
Db 2120 AGCAGGCCGCTGTCTTG 2103

Search completed: September 16, 2003, 22:44:45
Job time: 1307.45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 18:37:15 ; Search time 28.3366 Seconds
(without alignments)
535.803 Million cell updates/sec

Title: US-09-594-065-2

Perfect score: 18
Sequence: 1 agcagggcgcctgccttg 18

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 606776 seqs, 421745864 residues

Total number of hits satisfying chosen parameters: 1213552

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.4	91.1	1924	6	US-10-651-237-4
2	16.4	91.1	1924	6	US-10-651-237-7
3	15.4	85.6	201	7	US-60-495-114-38730
4	15.4	85.6	201	7	US-60-495-114-89612
5	15.4	85.6	201	7	US-60-495-114-89619
6	15.4	85.6	2441	6	US-10-425-114A-26835
7	15.4	85.6	2463	6	US-10-425-114A-26959
8	15.4	85.6	2470	7	US-60-490-890-1861
9	15.4	85.6	2487	6	US-10-425-114A-26285
10	15.4	85.6	2799	7	US-09-935-116A-2
11	15.4	85.6	26433	5	US-60-495-114-16387
12	15.4	85.6	45664	7	US-60-495-114-16957
13	15.4	85.6	384485	7	US-60-495-114-16382
14	15	83.3	1529	6	US-10-425-114A-27906
15	14.8	82.2	201	7	US-60-495-114-15241
16	14.8	82.2	201	7	US-60-495-114-54280
17	14.8	82.2	201	7	US-60-495-114-54313
18	14.8	82.2	201	7	US-60-495-114-68094
19	14.8	82.2	201	7	US-60-495-114-82382
20	14.8	82.2	201	7	US-60-495-114-82387
21	14.8	82.2	318	1	PCT-US03-11188-38
22	14.8	82.2	827	6	US-10-653-047-1247
23	14.8	82.2	924	6	US-10-425-114A-1273
24	14.8	82.2	1066	6	US-10-425-114A-33268
25	14.8	82.2	1135	6	US-10-425-114A-23292
26	14.8	82.2	1331	6	US-10-425-114A-27120

27	14.8	82.2	1373	6	US-10-425-114A-35800	Sequence 35800, A
28	14.8	82.2	1690	6	US-10-425-114A-33240	Sequence 33240, A
29	14.8	82.2	1759	7	US-60-495-114-1078	Sequence 1078, Ap
30	14.8	82.2	1806	6	US-10-425-114A-1679	Sequence 1679, Ap
31	14.8	82.2	1908	6	US-10-425-114A-21804	Sequence 21804, A
32	14.8	82.2	2054	6	US-10-425-114A-28366	Sequence 28366, A
33	14.8	82.2	2492	6	US-60-482-992-5	Sequence 31551, A
34	14.8	82.2	2520	7	US-10-425-114A-31551	Sequence 31551, A
35	14.8	82.2	4212	6	US-60-482-992-5	Sequence 26843, A
36	14.8	82.2	22806	7	US-60-495-114-16659	Sequence 16659, A
37	14.8	82.2	39902	7	US-60-487-610-20061	Sequence 20061, A
38	14.8	82.2	53795	1	PCT-US02-38582-124	Sequence 124, App
39	14.8	82.2	64607	7	US-60-495-114-16541	Sequence 16541, A
40	14.8	82.2	84010	7	US-60-487-610-19295	Sequence 19295, A
41	14.8	82.2	111982	7	US-60-487-610-19846	Sequence 19846, A
42	14.8	82.2	144434	7	US-60-495-114-16451	Sequence 16451, A
43	14.4	80.0	201	7	US-60-487-610-86638	Sequence 86638, A
44	14.4	80.0	423	6	US-10-084-846A-107	Sequence 107, App
45	14.4	80.0	1014	1	PCT-US03-23249-113	Sequence 113, App

ALIGNMENTS

RESULT 1
US-10-651-237-4/c
; Sequence 4, Application US/10651237
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: Colorectal Cancer Prognostics
; FILE REFERENCE: ADS-5003 US NP
; CURRENT APPLICATION NUMBER: US/10/651,237
; CURRENT FILING DATE: 2003-08-27
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1924
; TYPE: DNA
; ORGANISM: human
US-10-651-237-4

Query Match 91.1% Score 16.4; DB 6; Length 1924;
Best Local Similarity 94.4%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGGCCGCTGCTCTTG 18
DB 1412 AGCAGGCCGCTGCTCTTG 1395

RESULT 2
US-10-651-237-7/c
; Sequence 7, Application US/10651237
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: Colorectal Cancer Prognostics
; FILE REFERENCE: ADS-5003 US NP
; CURRENT APPLICATION NUMBER: US/10/651,237
; CURRENT FILING DATE: 2003-08-27
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1924
; TYPE: DNA
; ORGANISM: human
US-10-651-237-7

Query Match 91.1% Score 16.4; DB 6; Length 1924;
Best Local Similarity 94.4%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGCAGCCCGCTGTCCTTG 18
|||
DB 1412 AGCAGCCCGCTGTCCTTG 1395

RESULT 3

US-60-495-114-38730/C
Sequence 38730, Application US/60495114
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
FILE REFERENCE: CL001480
CURRENT APPLICATION NUMBER: US/60/495,114
CURRENT FILING DATE: 2003-08-15
NUMBER OF SEQ ID NOS: 91238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 38730
LENGTH: 201
TYPE: DNA
ORGANISM: Homo sapiens
US-60-495-114-38730

Query Match 85.6%; Score 15.4; DB 7; Length 201;
Best Local Similarity 94.1%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGCAGCCCGCTGTCCTT 17
|||
DB 92 AGCAGCCCGCTGTCCTT 76

RESULT 4

US-60-495-114-89612/C
Sequence 89612, Application US/60495114
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
FILE REFERENCE: CL001480
CURRENT APPLICATION NUMBER: US/60/495,114
CURRENT FILING DATE: 2003-08-15
NUMBER OF SEQ ID NOS: 91238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 89612
LENGTH: 201
TYPE: DNA
ORGANISM: Homo sapiens
US-60-495-114-89612

Query Match 85.6%; Score 15.4; DB 7; Length 201;
Best Local Similarity 94.1%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGCAGCCCGCTGTCCTT 17
|||
DB 64 AGCAGCCCGCTGTCCTT 48

RESULT 5

US-60-495-114-89619/C
Sequence 89619, Application US/60495114
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
FILE REFERENCE: CL001480
CURRENT APPLICATION NUMBER: US/60/495,114
CURRENT FILING DATE: 2003-08-15
NUMBER OF SEQ ID NOS: 91238

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 89619
LENGTH: 201
TYPE: DNA
ORGANISM: Homo sapiens
US-60-495-114-89619

Query Match 85.6%; Score 15.4; DB 7; Length 201;
Best Local Similarity 94.1%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGCAGCCCGCTGTCCTT 17
|||
DB 92 AGCAGCCCGCTGTCCTT 76

RESULT 6

US-10-425-114A-26835
Sequence 26835, Application US/10425114A
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114A
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 26835
LENGTH: 2441
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Clone ID: LIB4649-035-D9_FLI
US-10-425-114A-26835

Query Match 85.6%; Score 15.4; DB 6; Length 2441;
Best Local Similarity 94.1%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGCAGCCCGCTGTCCTT 17
|||
DB 49 AGCAGCCCGCTGTCCTT 65

RESULT 7

US-10-425-114A-26959
Sequence 26959, Application US/10425114A
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114A
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 26959
LENGTH: 2463
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Clone ID: LIB4676-059-G1_FLI
US-10-425-114A-26959

Query Match 85.6%; Score 15.4; DB 6; Length 2463;
Best Local Similarity 94.1%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 60 AGCAGGCCCTGTCTCTT 76

RESULT 8
US-60-490-890-1861
; Sequence 1861, Application US/60490890
; GENERAL INFORMATION:
; APPLICANT: Li, Marcha
; APPLICANT: Rupnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PSP
; CURRENT APPLICATION NUMBER: US/60/490,890
; CURRENT FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 2779
; SOFTWARE: Patent version 3.2
; SEQ ID NO 1861
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-490-890-1861

Query Match 85.6%; Score 15.4; DB 7; Length 2470;
Best Local Similarity 94.1%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGGCCCTGTCTCTT 17
|||
Db 78 AGCAGGCCCTGTCTCTT 94

RESULT 9
US-10-425-114A-26285
; Sequence 26285, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven B.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53333)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26285
; LENGTH: 2487
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; OTHER INFORMATION: Clone ID: LIB4119-174-G9_FLI
US-10-425-114A-26285

Query Match 85.6%; Score 15.4; DB 6; Length 2487;
Best Local Similarity 94.1%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGGCCCTGTCTCTT 17
|||
Db 95 AGCAGGCCCTGTCTCTT 111

RESULT 10

US-09-935-368A-2/c
; Sequence 2, Application US/09935368A
; GENERAL INFORMATION:
; APPLICANT: Cooper, Mark J.
; TITLE OF INVENTION: Expression System for Production of
; TITLE OF INVENTION: Therapeutic Proteins
; FILE REFERENCE: 003659.00010
; CURRENT APPLICATION NUMBER: US/09/935,368A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 09/473,646
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: PCT/US98/12777
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: US 60/050,356
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2799
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-368A-2

Query Match 85.6%; Score 15.4; DB 5; Length 2799;
Best Local Similarity 94.1%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGGCCCTGTCTCTT 17
|||
Db 332 AGCAGGCCCTGTCTCTT 316

RESULT 11
US-60-495-114-16387/c
; Sequence 16387, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO01480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16387
; LENGTH: 26433
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(26433)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-60-495-114-16387

Query Match 85.6%; Score 15.4; DB 7; Length 26433;
Best Local Similarity 94.1%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGGCCCTGTCTCTT 17
|||
Db 1366 AGCAGGCCCTGTCTCTT 1350

RESULT 12
US-60-495-114-16957/c
; Sequence 16957, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CLO01480

```

; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FASTSEQ for windows Version 4.0
; SEQ ID NO 16957
; LENGTH: 45664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(45664)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-60-495-114-16957

```

```

Query Match      85.6%; Score 15.4; DB 7; Length 45664;
Best Local Similarity 94.1%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 AGCAGCGCGCTGTCTT 17
      |||||
Db      37645 AGCAGCGCGCTGTCTT 37629

```

```

RESULT 13
US-60-495-114-16382/C
; Sequence 16382, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FASTSEQ for windows Version 4.0
; SEQ ID NO 16382
; LENGTH: 384485
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(384485)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-60-495-114-16382

```

```

Query Match      85.6%; Score 15.4; DB 7; Length 384485;
Best Local Similarity 94.1%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 AGCAGCGCGCTGTCTT 17
      |||||
Db      376152 AGCAGCGCGCTGTCTT 376136

```

```

RESULT 14
US-10-425-114A-27906
; Sequence 27906, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 27906
; LENGTH: 1529

```

```

; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4742-003-G8_FLI
US-10-425-114A-27906

```

```

Query Match      83.3%; Score 15; DB 6; Length 1529;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 GCAGCGCGCTGTCTT 16
      |||||
Db      127 GCAGCGCGCTGTCTT 141

```

```

RESULT 15
US-60-495-114-15241/C
; Sequence 15241, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FASTSEQ for windows Version 4.0
; SEQ ID NO 15241
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-495-114-15241

```

```

Query Match      82.2%; Score 14.8; DB 7; Length 201;
Best Local Similarity 88.9%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 AGCAGCGCGCTGTCTTG 18
      |||||
Db      65 AGCAGCGCGCTGTCTTG 48

```

```

Search completed: September 16, 2003, 22:47:53
Job time : 30.3366 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 18:28:15 ; Search time 2175.74 Seconds
(without alignments)
457.275 Million cell updates/sec

Title: US-09-594-065-3

Perfect score: 30
Sequence: 1 cccgcgtagtgcgtacgacctctgcaggg 30

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 33363688 seqs, 1658189874 residues

Total number of hits satisfying chosen parameters: 66727376

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_Main:*

```

1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.old.*
3: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US08_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US092A_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US092B_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US093A_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US093B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US098D_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US099D_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US099E_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US099F_COMB.seq.*

```

```

44: /cgn2_6/ptodata/2/pna/US100A_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US100B_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US101A_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US101B_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US102A_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US102B_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US103A_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US103B_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US104A_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US104B_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6004_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6023A_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6023B_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6026_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6027_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US6037_COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US6038_COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US6039_COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US6040_COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US6041_COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US6042_COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US6043_COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US6044_COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US6045_COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US6046_COMB.seq.*
102: /cgn2_6/ptodata/2/pna/US6047_COMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	25	US-09-594-065-3
2	23.8	79.3	2142	25	US-09-584-852-6658
3	20.4	68.0	714	51	US-10-369-493-41285
4	20.4	68.0	714	91	US-60-360-039-41285

5 20.4 68.0 2731748 49 US-10-297-465A-1 Sequence 1, Appl1
6 20.2 67.3 372 18 US-09-270-849B-77960 Sequence 77960, A
7 19.4 64.7 376 18 US-09-270-849B-51162 Sequence 51162, A
8 19.4 64.7 483 51 US-10-366-683-6704 Sequence 6704, Ap
9 19.4 64.7 480 18 US-10-419-128-6704 Sequence 6704, Ap
10 19.4 64.7 493 52 US-09-270-849B-169678 Sequence 169678, A
11 19.4 64.7 510 51 US-10-366-683-6599 Sequence 6599, Ap
12 19.4 64.7 510 52 US-10-419-128-6599 Sequence 6599, Ap
13 19.4 64.7 1302 49 US-10-283-122A-23788 Sequence 23788, A
14 19.4 64.7 1323 47 US-10-156-761-4750 Sequence 4750, Ap
15 19.4 64.7 1368 21 US-09-489-039A-6204 Sequence 6204, Ap
16 19.4 64.7 1368 53 US-10-446-203-6204 Sequence 6204, Ap
17 19.4 64.7 2022 49 US-10-263-929-79 Sequence 79, Appl
18 19.4 64.7 2022 46 US-10-140-771-8606 Sequence 8606, Ap
19 19.4 64.7 2079 91 US-10-360-207-8606 Sequence 8606, Ap
20 19.4 64.7 3084 51 US-10-366-683-6539 Sequence 6539, Ap
21 19.4 64.7 3084 51 US-10-419-128-6539 Sequence 6539, Ap
22 19.4 64.7 3273 51 US-10-366-683-6578 Sequence 6578, Ap
23 19.4 64.7 3273 52 US-10-419-128-6578 Sequence 6578, Ap
24 19.4 64.7 110079 1 PCT-US02-19457-96 Sequence 96, Appl
25 19.4 64.7 110079 47 US-10-175-523-96 Sequence 96, Appl
26 19.4 64.7 9025608 47 US-10-156-761-1 Sequence 1, Appl
27 19.2 64.0 106174 79 US-60-248-505-419 Sequence 419, Appl
28 19.2 63.3 482 18 US-09-270-849B-65208 Sequence 65208, A
29 19.2 63.3 487 22 US-09-521-640-194436 Sequence 194436, A
30 19.2 63.3 487 18 US-09-270-849B-43567 Sequence 43567, A
31 19.2 63.3 1485 53 US-10-437-963-102206 Sequence 102206, A
32 19.2 63.3 1485 53 US-10-438-246-11944 Sequence 11944, A
33 19.2 63.3 1485 53 US-10-438-246-21582 Sequence 21582, A
34 18.8 62.7 444 26 US-09-606-977-56997 Sequence 56997, A
35 18.8 62.7 444 68 US-60-141-233-56997 Sequence 56997, A
36 18.8 62.7 576 32 US-09-738-626-2147 Sequence 2147, Ap
37 18.8 62.7 578 72 US-60-182-316-4088 Sequence 4088, Ap
38 18.8 62.7 726 26 US-09-605-703B-1951 Sequence 1951, Ap
39 18.8 62.7 1368 16 US-09-191-989-3 Sequence 3, Appl
40 18.8 62.7 1593 47 US-10-156-761-5071 Sequence 5071, Ap
41 18.8 62.7 2936 16 US-09-191-989-1 Sequence 1, Appl
42 18.8 62.7 4454 16 US-09-191-989-2 Sequence 2, Appl
43 18.8 62.7 99916 34 US-09-816-095-3 Sequence 3, Appl
44 18.8 62.7 99916 53 US-10-634-905-3 Sequence 3, Appl
45 18.8 62.7 113462 22 US-09-528-237A-1860 Sequence 1860, Ap

ALIGNMENTS

RESULT 1
US-09-594-065-3
; Sequence 3, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: Walker-Peach, Cindy
; APPLICANT: DuBois, Dwight
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 an
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: IAC specific molecular beacon
US-09-594-065-3

Query Match 100.0%; Score 30; DB 25; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.011; Mismatches 0; Indels 0; Gaps 0;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CCTGCGTAGTGTACGACCTCTCGAGGG 30
|||||

Db 1 CCTGCGTAGTGTACGACCTCTCGAGGG 30

RESULT 2
US-09-584-852-6658
; Sequence 6658, Application US/09584852
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1118-001
; CURRENT APPLICATION NUMBER: US/09/584,852
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 60/135,618
; PRIOR FILING DATE: 1999-05-24
; PRIOR APPLICATION NUMBER: 60/135,711
; PRIOR FILING DATE: 1999-05-24
; PRIOR APPLICATION NUMBER: 60/135,627
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 8040
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6658
; LENGTH: 2142
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-584-852-6658

Query Match 79.3%; Score 23.8; DB 25; Length 2142;
Best Local Similarity 92.6%; Pred. No. 8.9;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 2 CCTGCGTAGTGTACGACCTCTCGAGG 28
446 CCTGCGTAGTGTACGACCTCTCGAGG 472

RESULT 3
US-10-369-493-41285/C
; Sequence 41285, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 41285
; LENGTH: 714
; TYPE: DNA
; ORGANISM: Xylella fastidiosa
US-10-369-493-41285

Query Match 68.0%; Score 20.4; DB 51; Length 714;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 1 CCTGCGTAGTGTACGACCTCTCGAGGG 30
403 CCTGCGTAGTGTACGACCTCTCGAGGG 374

RESULT 4
US-60-360-039-41285/C
; Sequence 41285, Application US/60360039
; GENERAL INFORMATION:

```

; APPLICANT: Cao, Yongwei
; ORGANISM: Xylella fastidiosa
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 41285
; LENGTH: 714
; TYPE: DNA
; ORGANISM: Xylella fastidiosa
US-60-360-039-41285

Query Match
Best Local Similarity 68.0%; Score 20.4; DB 91; Length 714;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTGCGTAGTGTACGACCTCTGCAGGG 30
DB 403 CCTGCGCATCGCTTCGACCTTGTGCAGGG 374

RESULT 5
US-10-297-465A-1
; Sequence 1, Application US/10297465A
; GENERAL INFORMATION:
; APPLICANT: Simpson, Andrew
; APPLICANT: Reinach, Fernando
; APPLICANT: Setubal, Joao
; APPLICANT: Medeiros, Joao
; APPLICANT: Aranda, Paulo
; TITLE OF INVENTION: Isolated Genome of Xylella fastidiosa and Uses Thereof
; FILE REFERENCE: PABSP 202 US (10213376)
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: PCT/IB01/01618
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/209,906
; PRIOR FILING DATE: 2001-06-17
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2731748
; TYPE: DNA
; ORGANISM: Xylella fastidiosa
US-10-297-465A-1

Query Match
Best Local Similarity 68.0%; Score 20.4; DB 49; Length 2731748;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTGCGTAGTGTACGACCTCTGCAGGG 30
DB 1446654 CCTGCGCATCGCTTCGACCTTGTGCAGGG 1446683

RESULT 6
US-09-270-849B-77960/C
; Sequence 77960, Application US/09270849B
; GENERAL INFORMATION:
; APPLICANT: Swimmer et al.
; TITLE OF INVENTION: Insect genome survey devices
; FILE REFERENCE:
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 195450
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 77960
; LENGTH: 372
```

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-77960

Query Match
Best Local Similarity 67.3%; Score 20.2; DB 18; Length 372;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TGGCGTAGTGTACGACCTCTGCAGG 28
DB 203 TGGCGTAGTGTACGACCTCTGCAGG 179

RESULT 7
US-09-270-849B-51162/C
; Sequence 51162, Application US/09270849B
; GENERAL INFORMATION:
; APPLICANT: Swimmer et al.
; TITLE OF INVENTION: Insect genome survey devices
; FILE REFERENCE:
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 195450
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51162
; LENGTH: 376
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-51162

Query Match
Best Local Similarity 64.7%; Score 19.4; DB 18; Length 376;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCTGCGTAGTGTACGACCTCTGCAGGG 30
DB 37 CCTGCGTAGTGTACGACCTCTCTGCAGGG 9

RESULT 8
US-10-366-683-6704/C
; Sequence 6704, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nolling, Jork
; APPLICANT: Deloushery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: PATH03-04
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6704
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-6704

Query Match
Best Local Similarity 64.7%; Score 19.4; DB 51; Length 483;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTGCGTAGTGTACGACCTCTGCAGG 29
DB 35 CCTGCGTAGTGTACGACCTCTGCAGG 7
```

```
RESULT 9
US-10-419-128-6704/c
; Sequence 6704, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; PRIOR FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6704
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-6704

Query Match      64.7%; Score 19.4; DB 52; Length 483;
Best Local Similarity 79.3%; Pred. No. 6.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTGCGTAGTGTGCTACGACCTCTCTGCAGG 29
DB 35 CCCGGCGCAGTGGCGGCGAGCGCTGCAGG 7

RESULT 10
US-09-270-849B-169678/c
; Sequence 169678, Application US/09270849B
; GENERAL INFORMATION:
; APPLICANT: Swimmer et al.
; TITLE OF INVENTION: Insect genome survey devices
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/270,849B
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 195450
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 169678
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-169678

Query Match      64.7%; Score 19.4; DB 18; Length 490;
Best Local Similarity 79.3%; Pred. No. 6.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCTGCGTAGTGTGCTACGACCTCTCTGCAGG 30
DB 244 CCTGCGTAGTGTGCTACGACCTCTATAGAG 216

RESULT 11
US-10-366-683-6599
; Sequence 6599, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nolling, Jork
; APPLICANT: Deloughery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: PATH03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
```

```
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6599
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-6599

Query Match      64.7%; Score 19.4; DB 51; Length 510;
Best Local Similarity 79.3%; Pred. No. 6.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTGCGTAGTGTGCTACGACCTCTCTGCAGG 29
DB 347 CCCGGCGCAGTGGCGGCGAGCGCTGCAGG 375

RESULT 12
US-10-419-128-6599
; Sequence 6599, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6599
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-6599

Query Match      64.7%; Score 19.4; DB 52; Length 510;
Best Local Similarity 79.3%; Pred. No. 6.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTGCGTAGTGTGCTACGACCTCTCTGCAGG 29
DB 347 CCCGGCGCAGTGGCGGCGAGCGCTGCAGG 375

RESULT 13
US-10-282-122A-23788
; Sequence 23788, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Travick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forey, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
```

PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent in version 3.1
SEQ ID NO 23788
LENGTH: 1302
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-10-282-122A-23788

Query Match 64.7% Score 19.4; DB 49; Length 1302;

Best Local Similarity 79.3%; Pred. No. 7.3e+02; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCTGCGTAGTGTGACGACCTCTGCAGG 29
Db 708 CCCGCGTGTGTGACGAGATTCTGCAGG 736

RESULT 14
US-10-156-761-4750
Sequence 4750, Application US/10156761
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 4750
LENGTH: 1323
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1323)
US-10-156-761-4750

Query Match 64.7% Score 19.4; DB 47; Length 1323;

Best Local Similarity 79.3%; Pred. No. 7.3e+02; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCTGCGTAGTGTGACGACCTCTGCAGG 29
Db 669 CCCGCGAAGCGGTGACGAGTTCTGCTGG 697

RESULT 15
US-09-489-039A-6204
Sequence 6204, Application US/09489039A
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709, 2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 6204
LENGTH: 1368
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6204

Query Match 64.7% Score 19.4; DB 21; Length 1368;

Best Local Similarity 79.3%; Pred. No. 7.4e+02; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCTGCGTAGTGTGACGACCTCTGCAGG 29
Db 771 CCCGCGTGTGTGACGAGATTCTGCAGG 799

Search completed: September 16, 2003, 22:44:55
Job time : 2185.74 secs

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 16, 2003, 18:37:15 ; Search time 47.2277 Seconds
(without alignments)
535.803 Million cell updates/sec

Title: US-09-594-065-3

Perfect score: 30

Sequence: 1 cccgcgcagtggtacgacctcgcaggg 30

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 606776 seqs, 421745864 residues

Total number of hits satisfying chosen parameters: 1213552

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_MA New.*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB_seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB_seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB_seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB_seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB_seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB_seq.*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB_seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20.4	68.0	2731748	6 US-10-297-465B-1	Sequence 1, Appl
2	18.8	62.7	4813087	5 US-09-947-914-75	Sequence 75, Appl
3	18	60.0	1627	6 US-10-425-114A-33391	Sequence 33391, A
4	17.8	59.3	1189	6 US-10-425-114A-3571	Sequence 3571, Ap
5	17.8	59.3	1268	6 US-10-425-114A-19153	Sequence 19153, A
6	17.8	59.3	1327	6 US-10-425-114A-5952	Sequence 5952, Ap
7	17.8	59.3	1746	6 US-10-425-114A-28420	Sequence 28420, A
8	17.6	58.7	1111	6 US-10-425-114A-1209	Sequence 1209, Ap
9	17.6	58.7	4829	6 US-10-428-487-41	Sequence 41, Appl
10	17.4	58.0	263	1 PCT-US03-18714-21709	Sequence 21709, A
11	17.4	58.0	734	7 US-60-480-590-5114	Sequence 5114, Ap
12	17.4	58.0	1445	6 US-10-350-923B-49	Sequence 49, Appl
13	17.4	58.0	1647	6 US-10-425-114A-3130	Sequence 3130, Ap
14	17.4	58.0	100537	7 US-60-485-450-11874	Sequence 11874, A
15	17.4	58.0	4813087	5 US-09-947-914-75	Sequence 75, Appl
16	17.2	57.3	201	7 US-60-487-610-31706	Sequence 31706, A
17	17.2	57.3	201	7 US-60-487-610-31706	Sequence 31706, A
18	17.2	57.3	201	7 US-60-487-610-31706	Sequence 31706, A
19	17.2	57.3	201	7 US-60-487-610-31706	Sequence 31706, A
20	17.2	57.3	201	7 US-60-487-610-31706	Sequence 31706, A
21	17.2	57.3	201	7 US-60-487-610-31706	Sequence 31706, A
22	17.2	57.3	201	7 US-60-487-610-31706	Sequence 31706, A
23	17.2	57.3	201	7 US-60-487-610-31706	Sequence 31706, A
24	17.2	57.3	201	7 US-60-487-610-31706	Sequence 31706, A
25	17.2	57.3	201	7 US-60-487-610-31706	Sequence 31706, A
26	17.2	57.3	201	7 US-60-487-610-31706	Sequence 31706, A

C	27	17.2	57.3	127313	7	US-60-487-610-19398	Sequence 19398, A
C	28	17	56.7	1224	6	US-10-296-115-154	Sequence 154, Appl
C	29	17	56.7	8059021	5	US-09-947-914-53	Sequence 53, Appl
C	30	16.8	56.0	201	7	US-60-487-610-57668	Sequence 57668, A
C	31	16.8	56.0	1981	7	US-60-487-610-793	Sequence 793, Appl
C	32	16.8	56.0	2017	6	US-10-296-115-464	Sequence 464, Appl
C	33	16.8	56.0	2052	7	US-60-493-007-2834	Sequence 2834, Ap
C	34	16.8	56.0	2189	5	US-09-908-576-7834	Sequence 7834, Ap
C	35	16.8	56.0	2189	7	US-60-487-610-789	Sequence 789, Appl
C	36	16.8	56.0	2189	7	US-60-487-610-789	Sequence 789, Appl
C	37	16.8	56.0	2320	7	US-60-487-610-788	Sequence 788, Appl
C	38	16.8	56.0	2366	7	US-60-487-610-794	Sequence 794, Appl
C	39	16.8	56.0	2473	7	US-60-487-610-795	Sequence 795, Appl
C	40	16.8	56.0	2657	7	US-60-487-610-791	Sequence 791, Appl
C	41	16.8	56.0	2663	7	US-60-487-610-792	Sequence 792, Appl
C	42	16.8	56.0	2715	7	US-60-487-610-796	Sequence 796, Appl
C	43	16.8	56.0	9041	7	US-60-487-196-332	Sequence 332, Appl
C	44	16.8	56.0	26562	7	US-60-487-610-20041	Sequence 20041, A
C	45	16.8	56.0	45317	7	US-60-487-610-19569	Sequence 19569, A

ALIGNMENTS

RESULT 1
US-10-297-465B-1
Sequence 1, Application US/10297465B
GENERAL INFORMATION:
APPLICANT: Simpson, Andrew
APPLICANT: Reinach, Fernando
APPLICANT: Setubal, Joao
APPLICANT: Medanis, Joao
TITLE OF INVENTION: Isolated Genome of *Xylella fastidiosa* and Uses Thereof
FILE REFERENCE: PABSP 202 US (10213376)
CURRENT FILING DATE: 2003-08-25
PRIOR APPLICATION NUMBER: US/10/297, 465B
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: PCT/IB01/01618
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: 60/209, 906
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 2731748
TYPE: DNA
ORGANISM: *Xylella fastidiosa*
US-10-297-465B-1
Query Match 68.0%; Score 20.4; DB 6; Length 2731748;
Beat Local Similarity 80.0%; Pred. No. 8.5;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 cccgcgcagtggtacgacctcgcaggg 30
DB 1446654 cccgcgcagtggtacgacctcgcaggg 1446683
RESULT 2
US-09-947-914-75
Sequence 75, Application US/09947914
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON CH
FILE REFERENCE: CLO01298
CURRENT FILING DATE: 2001-09-07
NUMBER OF SEQ ID NOS: 75
SEQ ID NO 75
LENGTH: 4813087
TYPE: DNA
ORGANISM: HUMAN

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(4813087)
OTHER INFORMATION: n = A,T,C or G
US-09-947-914-75

Query Match
Best Local Similarity 62.7%; Score 18.8; DB 5; Length 4813087;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCTCGCTAGTGTGTACGACCTCTCTGCAGG 30
DB 2035933 CCTGAGCATGTATACCGCTCTGCAGG 2035962

RESULT 3
US-10-425-114A-33391
Sequence 33391, Application US/10425114A
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114A
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 33391
LENGTH: 1627
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-2MFLMO17105A02_FLI
US-10-425-114A-33391

Query Match
Best Local Similarity 60.0%; Score 18; DB 6; Length 1627;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CTGCGTAGTGTGTACGACCTCTCTGCAG 28
DB 130 CGCGGAGGGGTTCCGATCTCTGCAG 155

RESULT 4
US-10-425-114A-3571/C
Sequence 3571, Application US/10425114A
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114A
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 3571
LENGTH: 1189
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700331949_FLI
US-10-425-114A-3571

Query Match 59.3%; Score 17.8; DB 6; Length 1189;

Best Local Similarity 75.9%; Pred. No. 87;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCTGCGTAGTGTGTACGACCTCTCTGCAGG 30
DB 1035 CCTATATAGTAGTATACCACTCTGCAG 1007

RESULT 5
US-10-425-114A-19153
Sequence 19153, Application US/10425114A
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114A
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 19153
LENGTH: 1268
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3137-002-D11_FLI
US-10-425-114A-19153

Query Match
Best Local Similarity 59.3%; Score 17.8; DB 6; Length 1268;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 GTGTACGACCTCTCTGCAGG 30
DB 207 GTGTACGACCTCTCTGCAGG 227

RESULT 6
US-10-425-114A-5952
Sequence 5952, Application US/10425114A
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114A
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 5952
LENGTH: 1327
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700549530_FLI
US-10-425-114A-5952

Query Match 59.3%; Score 17.8; DB 6; Length 1327;
Best Local Similarity 90.5%; Pred. No. 87;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 GTGTACGACCTCTCTGCAGG 30
DB 209 GTGTACGACCTCTCTGCAGG 229


```
RESULT 7
US-10-425-114A-28420/C
; Sequence 28420, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 28420
; LENGTH: 1746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4757-014-A3_FLI
US-10-425-114A-28420

Query Match          59.3%; Score 17.8; DB 6; Length 1746;
Best Local Similarity 75.9%; Pred. No. 89;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      2  CTTGGCTAGTGTAGACACTCTCTGCAGG 30
Db      1640 CCTATATAGTAGTACGACCACTCTGCAGG 1612

RESULT 8
US-10-425-114A-1209
; Sequence 1209, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 1209
; LENGTH: 1111
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700104758_FLI
US-10-425-114A-1209

Query Match          58.7%; Score 17.6; DB 6; Length 1111;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      6  CGTAGTGTAGACACTCTCTGCAGG 29
Db      846  CGCCGTGTAGACCACTCTGCAGG 869

RESULT 9
US-10-428-487-41
; Sequence 41, Application US/10428487
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA K.
; APPLICANT: GERBER, HANS-PETER
```

```
; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
; FILE REFERENCE: 09800080-0103
; CURRENT APPLICATION NUMBER: US/10/428,487
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/815,153
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,201
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-428-487-41

Query Match          58.7%; Score 17.6; DB 6; Length 4829;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      6  CGTAGTGTAGACACTCTCTGCAGG 29
Db      125  CTTGTGGGACCACTCTCTGCAGG 148

RESULT 10
PCT-US03-18714-21709/C
; Sequence 21709, Application PC/TUS0318714
; GENERAL INFORMATION:
; APPLICANT: Reguime Corporation
; TITLE OF INVENTION: Functional Sites
; FILE REFERENCE: 11207-029-228
; CURRENT APPLICATION NUMBER: PCT/US03/18714
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/387,887
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 60/387,910
; PRIOR FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 51999
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21709
; LENGTH: 263
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US03-18714-21709

Query Match          58.0%; Score 17.4; DB 1; Length 263;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      3  CTGCGTAGTGTAGACACTCTCTGCAGG 29
Db      151  CTGCGACTCGGAGGAACTCTCTGCAGG 125

RESULT 11
US-60-480-514/C
; Sequence 514, Application US/60480550
; GENERAL INFORMATION:
; APPLICANT: Monsanto
; APPLICANT: Monsanto
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Soybean Polymorphisms and Methods of Genotyping
; FILE REFERENCE: 38-15 (53382)
; CURRENT APPLICATION NUMBER: US/60/480,590
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 6578
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 514
; LENGTH: 734
; TYPE: DNA
; ORGANISM: Glycine max
US-60-480-514
```

Query Match 58.0%; Score 17.4; DB 7; Length 724;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 CCTGCGTAGTGATGACGACCTCTCTGCAG 28
|||||
Db 43 CCTGGGTAGTGGTTTGATCACCCTGAAG 17
|||||

RESULT 12

US-10-350-923B-49
; Sequence 49, Application US/10350923B
; GENERAL INFORMATION:
; APPLICANT: Dolson, Stanton B.
; APPLICANT: Ma, Xiao Jun
; TITLE OF INVENTION: Nucleic acid, polypeptides, vectors, and cells derived from active
; FILE REFERENCE: 503157-01
; CURRENT APPLICATION NUMBER: US/10/350,923B
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US/ 09/454,280
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: PCT/US99/28773
; PRIOR FILING DATE: 1999-06-12
; PRIOR APPLICATION NUMBER: US 60/111,006
; PRIOR FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 1445
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-350-923B-49

Query Match 58.0%; Score 17.4; DB 6; Length 1445;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 CCTGCGTAGTGATGACGACCTCTCTGCAG 28
|||||
Db 226 CCGAGTCTCTGTGACGACCTCTCTGCAG 252
|||||

RESULT 13

US-10-425-114A-3130/C
; Sequence 3130, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaka, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3130
; LENGTH: 1647
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700241792_FLI
US-10-425-114A-3130

Query Match 58.0%; Score 17.4; DB 6; Length 1647;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 CTGCGTAGTGATGACGACCTCTCTGCAG 29

Db 867 CTGCGAAGAGCTTCCCACTCTCTGCAG 841
|||||

RESULT 14

US-60-485-450-11874
; Sequence 11874, Application US/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; FILE REFERENCE: CLO01470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11874
; LENGTH: 100537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-485-450-11874

Query Match 58.0%; Score 17.4; DB 7; Length 100537;
Best Local Similarity 77.8%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 TCGGTAGTGATGACGACCTCTCTGCAGG 30
|||||
Db 6578 TGTGTAAGGTACTCTCTCTGCGG 6604
|||||

RESULT 15

US-09-947-914-75/C
; Sequence 75, Application US/09947914
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON CH
; FILE REFERENCE: CLO01298
; CURRENT APPLICATION NUMBER: US/09/947,914
; CURRENT FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 75
; SEQ ID NO 75
; LENGTH: 4813087
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(4813087)
; OTHER INFORMATION: n = A,T,C or G
US-09-947-914-75

Query Match 58.0%; Score 17.4; DB 5; Length 4813087;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCTGCGTAGTGATGACGACCTCTCTGCA 27
|||||
Db 1761296 CCTGCGTTGGGGAATATCTCTCTGCA 1761270
|||||

Search completed: September 16, 2003, 22:48:13
Job time : 67.2277 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 18:28:15 ; Search time 2393.32 Seconds
(without alignments)
457.275 Million cell updates/sec

Title: US-09-594-065-4

Perfect score: 33
Sequence: 1 cccctgcaactcgtctccctccagcatcgagg 33

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3363688 seqs, 1658189874 residues

Total number of hits satisfying chosen parameters: 66727376

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Pending Parents_MA Main:*

```

1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US080_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US092_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US092B_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US093A_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US093B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US098D_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US099E_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US099F_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US099F_COMB.seq.*

```

```

44: /cgn2_6/ptodata/2/pna/US100A_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US100B_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US101A_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US101B_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US102A_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US102B_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US103A_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US103B_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US104A_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US104B_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6004_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6023A_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6023B_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6026_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6027_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US6037_COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US6038_COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US6039_COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US6040_COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US6041_COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US6042_COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US6043_COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US6044_COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US6045_COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US6046_COMB.seq.*
102: /cgn2_6/ptodata/2/pna/US6047_COMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32.6	98.8	33	US-09-594-065-4	Sequence 4, Appli
2	32.6	98.8	33	US-09-594-065-10	Sequence 10, Appli
3	31.4	95.2	33	US-09-594-065-11	Sequence 11, Appli
4	22.6	68.5	39	US-09-594-065-7	Sequence 7, Appli

```

c 5 22.6 68.5 109 25 US-09-594-065-13 Sequence 13, Appl
6 22.6 68.5 2645 6 US-08-123-456-141 Sequence 141, App
7 22.6 68.5 2645 18 US-09-297-477A-141 Sequence 141, App
8 22.6 68.5 2645 43 US-09-994-404-141 Sequence 141, App
c 9 22.6 68.5 2841 6 US-08-123-456-78 Sequence 78, Appl
c 10 22.6 68.5 2841 18 US-09-297-477A-78 Sequence 78, Appl
c 11 22.6 68.5 2841 43 US-09-994-404-78 Sequence 78, Appl
c 12 22.6 68.5 117213 6 US-08-123-456-217 Sequence 217, App
c 13 22.6 68.5 117213 18 US-09-297-477A-217 Sequence 217, App
c 14 22.6 68.5 117213 43 US-09-994-404-217 Sequence 217, App
15 22.6 68.5 154746 1 PCT-US01-11372-8 Sequence 8, Appl
16 22.6 68.5 154746 2 PCT-US01-11372-8 Sequence 8, Appl
17 22.6 68.5 154746 34 US-09-827-668-8 Sequence 8, Appl
18 22.6 68.5 154746 15 US-09-065-511-3362 Sequence 3362, Ap
19 22.6 67.9 188 23 US-09-539-331D-25769 Sequence 25769, A
c 20 22.4 67.9 201 100 US-60-452-680-72811 Sequence 72811, A
c 21 22.4 67.9 247 14 US-08-937-141-1744 Sequence 1744, Ap
c 22 22.4 67.9 247 24 US-09-540-229-75778 Sequence 75778, A
c 23 22.4 67.9 247 56 US-60-027-778-1744 Sequence 1744, Ap
c 24 22.4 67.9 257 15 US-09-076-667-1489 Sequence 1489, Ap
c 25 22.4 67.9 257 24 US-09-540-229-89186 Sequence 89186, A
c 26 22.4 67.9 257 58 US-60-048-002-1489 Sequence 1489, Ap
c 27 22.4 67.9 259 24 US-09-540-229-100072 Sequence 100072,
28 22.4 67.9 271 16 US-09-105-427-127 Sequence 127, App
29 22.4 67.9 271 24 US-09-540-208-1734 Sequence 1734, Ap
c 30 22.4 67.9 294 24 US-09-540-764-49141 Sequence 49141, A
c 31 22.4 67.9 294 50 US-10-349-781-49141 Sequence 49141, A
c 32 22.4 67.9 319 15 US-09-079-506-42 Sequence 42, Appl
c 33 22.4 67.9 319 24 US-09-540-208-5526 Sequence 5526, Ap
c 34 22.4 67.9 478 22 US-09-528-409-21764 Sequence 21764, A
c 35 22.4 67.9 478 39 US-09-933-524-21764 Sequence 21764, A
c 36 22.4 67.9 478 39 US-09-933-524-21764 Sequence 21764, A
c 37 22.4 67.9 487 18 US-09-277-227-18868 Sequence 18868, A
c 38 22.4 67.9 487 19 US-09-346-956-16027 Sequence 16027, A
c 39 22.4 67.9 487 38 US-09-904-703-16027 Sequence 16027, A
c 40 22.4 67.9 487 38 US-09-904-627-18868 Sequence 18868, A
c 41 22.4 67.9 502 44 US-10-023-386-11498 Sequence 11498, A
c 42 22.4 67.9 551 44 US-10-023-386-25205 Sequence 25205, A
c 43 22.4 67.9 573 24 US-09-540-212A-61467 Sequence 61467, A
c 44 22.4 67.9 601 40 US-09-947-907-2022 Sequence 2022, Ap
c 45 22.4 67.9 601 40 US-09-949-016-37689 Sequence 37689, A

```

ALIGNMENTS

```

RESULT 1
US-09-594-065-4
; Sequence 4, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: Walkerpetch, Cindy
; APPLICANT: Dubois, Dwight
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 an
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HSV-type common (TC) molecular beacon
US-09-594-065-4

```

```

Query Match 98.8%; Score 32.6; DB 25; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CCCTGCAAACTGTCGTCTCTCCAGCATGCAGG 33
|||||

```

```

Db 1 CCCTGCAAACTGTCGTCTCTCCAGCATGCAGG 33
RESULT 2
US-09-594-065-10
; Sequence 10, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: Walkerpetch, Cindy
; APPLICANT: Dubois, Dwight
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 an
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HSV type-1 specific molecular beacon
US-09-594-065-10

```

```

Query Match 98.8%; Score 32.6; DB 25; Length 33;
Best Local Similarity 97.0%; Pred. No. 0.0083;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
1 CCCTGCAAACTGTCGTCTCTCCAGCATGCAGG 33
|||||

```

```

RESULT 3
US-09-594-065-11
; Sequence 11, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: Walkerpetch, Cindy
; APPLICANT: Dubois, Dwight
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 an
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HSV type-2 specific molecular beacon
US-09-594-065-11

```

```

Query Match 95.2%; Score 31.4; DB 25; Length 33;
Best Local Similarity 97.0%; Pred. No. 0.026;
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 CCCTGCAAACTGTCGTCTCTCCAGCATGCAGG 33
|||||

```

```

Db 1 CCCTGCAAACTGTCGTCTCTCCAGCATGCAGG 33
RESULT 4
US-09-594-065-7/c
; Sequence 7, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: Walkerpetch, Cindy
; APPLICANT: Dubois, Dwight
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 an
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
1 CCCTGCAAACTGTCGTCTCTCCAGCATGCAGG 33
|||||

```

```
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Single stranded oligonucleotide molecular beacon target for HSV 1
US-09-594-065-7
```

```
Query Match          68.5%; Score 22.6; DB 25; Length 39;
Best Local Similarity 95.7%; Pred. No. 1.1e+02;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      6 CAAACTCGTGTCTCCGAGCATG 28
Db      31 CAAACTCGTGTCTCCGAGCATG 9
```

```
RESULT 5
US-09-594-065-13/c
; Sequence 13, Application US/09594065
; GENERAL INFORMATION:
; APPLICANT: WalkerPeach, Cindy
; APPLICANT: Dubois, Dwight
; TITLE OF INVENTION: Compositions, methods and kits for Herpes Simplex Virus Type 1 an
; FILE REFERENCE: 25436/1280
; CURRENT APPLICATION NUMBER: US/09/594,065
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 109
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence amplified by PCR for HSV type-2
US-09-594-065-13
```

```
Query Match          68.5%; Score 22.6; DB 25; Length 109;
Best Local Similarity 95.7%; Pred. No. 1.2e+02;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      6 CAAACTCGTGTCTCCGAGCATG 28
Db      60 CAAACTCGTGTCTCCGAGCATG 38
```

```
RESULT 6
US-08-123-456-141
; Sequence 141, Application US/08123456
; GENERAL INFORMATION:
; APPLICANT: ESSER, KLAUS M.
; APPLICANT: CHAN, JOHN Y.
; APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
; APPLICANT: DELVECCHIO, ALFRED MICHAEL
; APPLICANT: DILLON, SUSAN B.
; APPLICANT: LEARY, JEFFREY JOSEPH
; APPLICANT: SUTTON, DAVID
; TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
; FILE REFERENCE: P50583
; CURRENT APPLICATION NUMBER: US/08/123,456
; CURRENT FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: US 60/030,279
; PRIOR FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: US 60/049,018
; PRIOR FILING DATE: 1997-06-09
; NUMBER OF SEQ ID NOS: 303
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 141
```

```
; LENGTH: 2645
; TYPE: DNA
; ORGANISM: Herpes simplex
US-08-123-456-141
```

```
Query Match          68.5%; Score 22.6; DB 6; Length 2645;
Best Local Similarity 95.7%; Pred. No. 1.8e+02;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      6 CAAACTCGTGTCTCCGAGCATG 28
Db      755 CAAACTCGTGTCTCCGAGCATG 777
```

```
RESULT 7
US-09-297-477A-141
; Sequence 141, Application US/09297477A
; GENERAL INFORMATION:
; APPLICANT: ESSER, KLAUS M.
; APPLICANT: CHAN, JOHN Y.
; APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
; APPLICANT: DELVECCHIO, ALFRED MICHAEL
; APPLICANT: DILLON, SUSAN B.
; APPLICANT: LEARY, JEFFREY JOSEPH
; APPLICANT: SUTTON, DAVID
; TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
; FILE REFERENCE: P50583
; CURRENT APPLICATION NUMBER: US/09/297,477A
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 60/030,279
; PRIOR FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: US 60/049,018
; PRIOR FILING DATE: 1997-06-09
; NUMBER OF SEQ ID NOS: 303
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 141
; LENGTH: 2645
; TYPE: DNA
; ORGANISM: Herpes simplex
US-09-297-477A-141
```

```
Query Match          68.5%; Score 22.6; DB 18; Length 2645;
Best Local Similarity 95.7%; Pred. No. 1.8e+02;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      6 CAAACTCGTGTCTCCGAGCATG 28
Db      755 CAAACTCGTGTCTCCGAGCATG 777
```

```
RESULT 8
US-09-994-404-141
; Sequence 141, Application US/09994404
; GENERAL INFORMATION:
; APPLICANT: ESSER, KLAUS M.
; APPLICANT: CHAN, JOHN Y.
; APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
; APPLICANT: DELVECCHIO, ALFRED MICHAEL
; APPLICANT: DILLON, SUSAN B.
; APPLICANT: LEARY, JEFFREY JOSEPH
; APPLICANT: SUTTON, DAVID
; TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
; FILE REFERENCE: P50583
; CURRENT APPLICATION NUMBER: US/09/994,404
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/297,477
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 60/030,279
; PRIOR FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: US 60/049,018
; PRIOR FILING DATE: 1997-06-09
```

NUMBER OF SEQ ID NOS: 303
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 141
LENGTH: 2645
TYPE: DNA
ORGANISM: Herpes simplex
US-09-994-404-141

Query Match
Best Local Similarity 68.5%; Score 22.6; DB 43; Length 2645;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CAACCTCGTGTCTCTCCAGCATG 28
Db 755 CAACCTCGTGTCTCTCCAGCATG 777

RESULT 9
US-08-123-456-78/c
Sequence 78, Application US/08123456
GENERAL INFORMATION:
APPLICANT: ESSER, KLAUS M.
APPLICANT: CHAN, JOHN Y.
APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
APPLICANT: DELVECCHIO, ALFRED MICHAEL
APPLICANT: DILON, SUSAN B.
APPLICANT: LEARY, JEFFREY JOSEPH
APPLICANT: SUTTON, DAVID
TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
FILE REFERENCE: P50583
CURRENT APPLICATION NUMBER: US/08/123,456
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: US 60/030,279
PRIOR FILING DATE: 1999-11-04
PRIOR APPLICATION NUMBER: US 60/049,018
PRIOR FILING DATE: 1997-06-09
NUMBER OF SEQ ID NOS: 303
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 78
LENGTH: 2841
TYPE: DNA
ORGANISM: Herpes simplex
NAME/KEY: unknown
LOCATION: (2591) (2825) (2833)
OTHER INFORMATION:
US-08-123-456-78

Query Match
Best Local Similarity 68.5%; Score 22.6; DB 6; Length 2841;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CAACCTCGTGTCTCTCCAGCATG 28
Db 2071 CAACCTCGTGTCTCTCCAGCATG 2049

RESULT 10
US-09-297-477A-78/c
Sequence 78, Application US/09297477A
GENERAL INFORMATION:
APPLICANT: ESSER, KLAUS M.
APPLICANT: CHAN, JOHN Y.
APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
APPLICANT: DELVECCHIO, ALFRED MICHAEL
APPLICANT: DILON, SUSAN B.
APPLICANT: LEARY, JEFFREY JOSEPH
APPLICANT: SUTTON, DAVID
TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
FILE REFERENCE: P50583
CURRENT APPLICATION NUMBER: US/09/297,477A

CURRENT FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US 60/030,279
PRIOR FILING DATE: 1999-11-04
PRIOR APPLICATION NUMBER: US 60/049,018
PRIOR FILING DATE: 1997-06-09
NUMBER OF SEQ ID NOS: 303
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 78
LENGTH: 2841
TYPE: DNA
ORGANISM: Herpes simplex
NAME/KEY: unsure
LOCATION: (2591) (2825) (2833)
OTHER INFORMATION: where n can equal A,G,C,T/U
US-09-297-477A-78

Query Match
Best Local Similarity 68.5%; Score 22.6; DB 18; Length 2841;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CAACCTCGTGTCTCTCCAGCATG 28
Db 2071 CAACCTCGTGTCTCTCCAGCATG 2049

RESULT 11
US-09-994-404-78/c
Sequence 78, Application US/09994404
GENERAL INFORMATION:
APPLICANT: ESSER, KLAUS M.
APPLICANT: CHAN, JOHN Y.
APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
APPLICANT: DELVECCHIO, ALFRED MICHAEL
APPLICANT: DILON, SUSAN B.
APPLICANT: LEARY, JEFFREY JOSEPH
APPLICANT: SUTTON, DAVID
TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
FILE REFERENCE: P50583
CURRENT APPLICATION NUMBER: US/09/994,404
PRIOR FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/297,477
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US 60/030,279
PRIOR FILING DATE: 1999-11-04
PRIOR APPLICATION NUMBER: US 60/049,018
PRIOR FILING DATE: 1997-06-09
NUMBER OF SEQ ID NOS: 303
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 78
LENGTH: 2841
TYPE: DNA
ORGANISM: Herpes simplex
NAME/KEY: unsure
LOCATION: (2591) (2825) (2833)
OTHER INFORMATION: where n can equal A,G,C,T/U
US-09-994-404-78

Query Match
Best Local Similarity 68.5%; Score 22.6; DB 43; Length 2841;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CAACCTCGTGTCTCTCCAGCATG 28
Db 2071 CAACCTCGTGTCTCTCCAGCATG 2049

RESULT 12
US-08-123-456-217
Sequence 217, Application US/08123456
GENERAL INFORMATION:

APPLICANT: ESSER, KLAUS M.
APPLICANT: CHAN, JOHN Y.
APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
APPLICANT: DELVECCHIO, ALFRED MICHAEL
APPLICANT: DILLON, SUSAN B.
APPLICANT: LEARY, JEFFREY JOSEPH
APPLICANT: SUTTON, DAVID
TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
FILE REFERENCE: P50583
CURRENT APPLICATION NUMBER: US/08/123,456
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: US 60/030,279
PRIOR FILING DATE: 1999-11-04
PRIOR APPLICATION NUMBER: US 60/049,018
PRIOR FILING DATE: 1997-06-09
NUMBER OF SEQ ID NOS: 303
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 217
LENGTH: 117213
TYPE: DNA
ORGANISM: Herpes simplex
FEATURE:
NAME/KEY: unknown
LOCATION: (100831) (100832) (100833) (111047) (111050) (111051)
OTHER INFORMATION:
US-08-123-456-217

Query Match 68.5%; Score 22.6; DB 6; Length 117213;
Best Local Similarity 95.7%; Pred. No. 2.8e+02;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 CAAACTGCTGCTCTCCAGCATG 28
DB 45257 CAAACTGCTGCTCTCCAGCATG 45279

RESULT 13
US-09-297-477A-217
Sequence 217, Application US/09297477A
GENERAL INFORMATION:
APPLICANT: ESSER, KLAUS M.
APPLICANT: CHAN, JOHN Y.
APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
APPLICANT: DELVECCHIO, ALFRED MICHAEL
APPLICANT: DILLON, SUSAN B.
APPLICANT: LEARY, JEFFREY JOSEPH
APPLICANT: SUTTON, DAVID
TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
FILE REFERENCE: P50583
CURRENT APPLICATION NUMBER: US/09/297,477A
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US 60/030,279
PRIOR FILING DATE: 1999-11-04
PRIOR APPLICATION NUMBER: US 60/049,018
PRIOR FILING DATE: 1997-06-09
NUMBER OF SEQ ID NOS: 303
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 217
LENGTH: 117213
TYPE: DNA
ORGANISM: Herpes simplex
FEATURE:
NAME/KEY: unsure
LOCATION: (100831) (100832) (100833) (111047) (111050) (111051)
OTHER INFORMATION: where n can equal A,G,C,T/U
US-09-297-477A-217

Query Match 68.5%; Score 22.6; DB 18; Length 117213;
Best Local Similarity 95.7%; Pred. No. 2.8e+02;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 CAAACTGCTGCTCTCCAGCATG 28
DB 45257 CAAACTGCTGCTCTCCAGCATG 45279

RESULT 14
US-09-994-404-217
Sequence 217, Application US/09994404
GENERAL INFORMATION:
APPLICANT: ESSER, KLAUS M.
APPLICANT: CHAN, JOHN Y.
APPLICANT: DABROWSKI-AMARAL, CHRISTINE ELLEN
APPLICANT: DELVECCHIO, ALFRED MICHAEL
APPLICANT: DILLON, SUSAN B.
APPLICANT: LEARY, JEFFREY JOSEPH
APPLICANT: SUTTON, DAVID
TITLE OF INVENTION: NOVEL CODING SEQUENCES FROM HERPES
FILE REFERENCE: P50583
CURRENT APPLICATION NUMBER: US/09/994,404
PRIOR FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/297,477
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US 60/030,279
PRIOR FILING DATE: 1999-11-04
PRIOR APPLICATION NUMBER: US 60/049,018
PRIOR FILING DATE: 1997-06-09
NUMBER OF SEQ ID NOS: 303
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 217
LENGTH: 117213
TYPE: DNA
ORGANISM: Herpes simplex
FEATURE:
NAME/KEY: unsure
LOCATION: (100831) (100832) (100833) (111047) (111050) (111051)
OTHER INFORMATION: where n can equal A,G,C,T/U
US-09-994-404-217

Query Match 68.5%; Score 22.6; DB 43; Length 117213;
Best Local Similarity 95.7%; Pred. No. 2.8e+02;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 CAAACTGCTGCTCTCCAGCATG 28
DB 45257 CAAACTGCTGCTCTCCAGCATG 45279

RESULT 15
PCT-US01-11372-8
Sequence 8, Application PC/TUS0111372
GENERAL INFORMATION:
APPLICANT: ORSON, FRANK
APPLICANT: KINSEY, BERNA
APPLICANT: BHOAGAL, BALBIR
TITLE OF INVENTION: MACROAGGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION
FILE REFERENCE: P01949051/10004014
CURRENT APPLICATION NUMBER: PCT/US01/11372
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/195,680
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 154746
TYPE: DNA
ORGANISM: HERPESVIRUS 2
PCT-US01-11372-8

Query Match 68.5%; Score 22.6; DB 1; Length 154746;
Best Local Similarity 95.7%; Pred. No. 2.9e+02;

Matches	22;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	6	CAACTCGTGTCTCTCCAGCATG	28						
Db	54077	CAACTCGTGTCTCTCCAGCATG	54099						

Search completed: September 16, 2003, 22:44:57
Job time : 2395.32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 16, 2003, 18:37:15 ; Search time 51.9505 Seconds
(without alignment)
535.803 Million cell updates/sec

Title: US-09-594-065-4

Perfect score: 33

Sequence: 1 cccgtcaactcgtgkctccacgacgacgag 33

Scoring table: IDENTITY NUC
Gapop 10'-0', Gapext 1.0

Searched: 606776 seqs, 421745864 residues

Total number of hits satisfying chosen parameters: 1213552

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_New: +
1: /cgn2_6/ptodata/2/pna/PCT_US03-21510-30
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq: +
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq: +
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq: +
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq: +
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq: +
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq: +

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20.2	61.2	2893	7	US-60-496-393-1 Sequence 1, Appl
2	20.2	61.2	3059	1	PCT-US03-21510-30 Sequence 30, Appl
3	20.2	61.2	3233	7	US-60-490-890-2277 Sequence 2277, Ap
4	20	60.6	36377	7	US-60-487-610-19291 Sequence 19291, A
5	20	60.6	46883	7	US-60-487-610-19887 Sequence 19887, A
6	20	60.6	4813087	5	US-09-947-914-75 Sequence 75, Appl
7	19.8	60.0	817	7	US-60-493-007-3124 Sequence 3124, Ap
8	19.8	60.0	904	7	US-60-487-610-1136 Sequence 1136, Ap
9	19.8	60.0	3679	5	US-09-908-576-244 Sequence 244, Ap
10	19.8	60.0	26000	7	US-60-495-114-16801 Sequence 16801, A
11	19.8	60.0	29102	7	US-60-487-610-19825 Sequence 19825, A
12	19.8	60.0	29102	7	US-60-485-450-12335 Sequence 12335, A
13	19.8	60.0	60476	7	US-60-487-610-19824 Sequence 19824, A
14	19.8	60.0	60476	7	US-60-485-450-12207 Sequence 12207, A
15	19.8	60.0	62312	7	US-60-495-114-16591 Sequence 16591, A
16	19.8	57.6	1801	6	US-10-425-114A-30109 Sequence 30109, A
17	19.8	57.6	513030	7	US-60-487-610-19737 Sequence 19737, A
18	18.8	57.0	3609	7	US-60-493-007-3356 Sequence 3356, Ap
19	18.8	57.0	4749	1	PCT-US03-073408-3 Sequence 3, Appl
20	18.8	57.0	4974	1	PCT-US03-073408-10 Sequence 10, Appl
21	18.8	57.0	43782	7	US-60-485-450-12181 Sequence 12181, A
22	18.8	57.0	47493	1	PCT-US02-36071A-55 Sequence 55, Appl
23	18.6	56.4	2600	6	US-10-425-114A-34704 Sequence 34704, A
24	18.6	56.4	83080	5	US-09-897-516A-423 Sequence 423, Ap
25	18.6	56.4	83080	5	US-09-897-516A-3852 Sequence 3852, Ap
26	18.6	56.4	83080	5	US-09-897-516A-3853 Sequence 3853, Ap

27	18.6	56.4	83080	5	US-09-897-516A-3854	Sequence 3854, Ap
28	18.6	56.4	83080	5	US-09-897-516A-3855	Sequence 3855, Ap
29	18.6	56.4	83080	5	US-09-897-516A-3856	Sequence 3856, Ap
30	18.6	56.4	83080	5	US-09-897-516A-3857	Sequence 3857, Ap
31	18.6	56.4	83080	5	US-09-897-516A-3858	Sequence 3858, Ap
32	18.6	56.4	83080	5	US-09-897-516A-3859	Sequence 3859, Ap
33	18.6	56.4	83080	5	US-09-897-516A-3860	Sequence 3860, Ap
34	18.6	56.4	83080	5	US-09-897-516A-3861	Sequence 3861, Ap
35	18.6	56.4	83080	5	US-09-897-516A-3862	Sequence 3862, Ap
36	18.6	56.4	83080	5	US-09-897-516A-3863	Sequence 3863, Ap
37	18.6	56.4	83080	5	US-09-897-516A-3864	Sequence 3864, Ap
38	18.6	56.4	83080	5	US-09-897-516A-3865	Sequence 3865, Ap
39	18.6	56.4	83080	5	US-09-897-516A-3866	Sequence 3866, Ap
40	18.6	56.4	83080	5	US-09-897-516A-3867	Sequence 3867, Ap
41	18.6	56.4	83080	5	US-09-897-516A-3868	Sequence 3868, Ap
42	18.6	56.4	83080	5	US-09-897-516A-3869	Sequence 3869, Ap
43	18.6	56.4	83080	5	US-09-897-516A-3870	Sequence 3870, Ap
44	18.6	56.4	83080	5	US-09-897-516A-3871	Sequence 3871, Ap
45	18.6	56.4	83080	5	US-09-897-516A-3872	Sequence 3872, Ap

ALIGNMENTS

```

RESULT 1
US-60-496-393-1/c
; Sequence 1, Application US/60496393
; GENERAL INFORMATION:
; APPLICANT: Amit, Ido
; APPLICANT: Yarkir, Ilat
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES AND ANTIBODIES AND USE THEREOF IN
; TITLE OF INVENTION: TREATING TSG101-ASSOCIATED DISEASES
; FILE REFERENCE: 25725
; CURRENT APPLICATION NUMBER: US/60/496,393
; CURRENT FILING DATE: 2003-08-20
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-496-393-1

Query Match      61.2%  Score 20.2;  DB 7;  Length 2893;
Best Local Similarity 81.5%; Pred. No. 24;
Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy      1  CCCTGCAACTCGTGTCTCCACGACAT 27
Db      902 CCCTGCAACTCGTGTCTCCACGACAT 876

RESULT 2
PCT-US03-21510-30
; Sequence 30, Application PC/TUS0321510
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: NP215 AS MODIFIERS OF THE p21 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-046C-PC
; CURRENT APPLICATION NUMBER: PCT/US03/21510
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US 60/394,795
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/401,739
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/411,010
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/437,158
; PRIOR FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30

```

LENGTH: 3059
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US03-21510-30

Query Match 61.2% Score 20.2; DB 1; Length 3059;
Best Local Similarity 75.8% Pred. No. 24;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1 CCTGCAAACTGCTGCTCTCCAGCATGACGG 33
797 CCTGCAAACTGCTGCTCTCCAGCCCGGGG 829

RESULT 3
US-60-490-890-2277
Sequence 2277, Application US/60490890
GENERAL INFORMATION:
APPLICANT: Li, Martha
APPLICANT: Rupnow, Brent A.
APPLICANT: Webster, Kevin R.
APPLICANT: Jackson, Donald
APPLICANT: Mong, Tai W.
TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
FILE REFERENCE: D0310 PSP
CURRENT APPLICATION NUMBER: US/60/490,890
CURRENT FILING DATE: 2003-07-29
NUMBER OF SEQ ID NOS: 2779
SOFTWARE: FastSeq for Windows Version 3.2
SEQ ID NO 2277
LENGTH: 3233
TYPE: DNA
ORGANISM: Homo sapiens
US-60-490-890-2277

Query Match 61.2% Score 20.2; DB 7; Length 3233;
Best Local Similarity 75.8% Pred. No. 24;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1 CCTGCAAACTGCTGCTCTCCAGCATGACGG 33
786 CCTGCAAACTGCTGCTCTCCAGCCCGGG 818

RESULT 4
US-60-487-610-19291/C
Sequence 19291, Application US/60487610
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: HUANG, Hongjin
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001469
CURRENT APPLICATION NUMBER: US/60/487,610
CURRENT FILING DATE: 2003-07-17
NUMBER OF SEQ ID NOS: 97101
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19291
LENGTH: 36377
TYPE: DNA
ORGANISM: Homo sapiens
US-60-487-610-19291

Query Match 60.6% Score 20; DB 7; Length 36377;
Best Local Similarity 76.7% Pred. No. 43;
Matches 23; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

2 CCTGCAAACTGCTGCTCTCCAGCATGACG 31
5011 CCTGCAAGGCTGTGCTCTCCAGGACG 4982

RESULT 5
US-60-487-610-19887
Sequence 19887, Application US/60487610
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: HUANG, Hongjin
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001469
CURRENT APPLICATION NUMBER: US/60/487,610
CURRENT FILING DATE: 2003-07-17
NUMBER OF SEQ ID NOS: 97101
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19887
LENGTH: 46883
TYPE: DNA
ORGANISM: Homo sapiens
US-60-487-610-19887

Query Match 60.6% Score 20; DB 7; Length 46883;
Best Local Similarity 76.7% Pred. No. 45;
Matches 23; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

1 CCTGCAAACTGCTGCTCTCCAGCATGCA 30
14028 CCTGCAACCTGGGTAGCTCCAGCATGCA 14057

RESULT 6
US-09-947-914-75/C
Sequence 75, Application US/09947914
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig
TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON CH
TITLE OF INVENTION: 21, METHODS OF DETECTION, AND USES THEREOF
FILE REFERENCE: CL001298
CURRENT APPLICATION NUMBER: US/09/947,914
CURRENT FILING DATE: 2001-09-07
NUMBER OF SEQ ID NOS: 75
SEQ ID NO 75
LENGTH: 4813087
TYPE: DNA
ORGANISM: HUMAN
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(4813087)
OTHER INFORMATION: n = A,T,C or G
US-09-947-914-75

Query Match 60.6% Score 20; DB 5; Length 4813087;
Best Local Similarity 76.7% Pred. No. 74;
Matches 23; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

1 CCTGCAAACTGCTGCTCTCCAGCATGCA 30
3040475 CCTGCAACCTGGGTAGCTCCAGCATGCA 3040446

RESULT 7
US-60-493-007-3124/C
Sequence 3124, Application US/60493007
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Higgs, Brandon
APPLICANT: Casle, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Primary rat hepatocyte toxicity modeling
FILE REFERENCE: 44921-5113-01-US
CURRENT APPLICATION NUMBER: US/60/493,007
CURRENT FILING DATE: 2003-08-07

```

; PRIOR APPLICATION NUMBER: US 60/353,171
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/363,534
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/371,135
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/371,134
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/370,248
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/371,150
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/371,413
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: US 60/373,601
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/374,139
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: US 60/394,253
; PRIOR FILING DATE: 2002-07-09
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 3518
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3124
; LENGTH: 817
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_019194
US-60-493-007-3124

```

```

Query Match          60.0%; Score 19.8; DB 7; Length 817;
Best Local Similarity 84.0%; Pred. No. 29;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy      3 CTGCAACTCGTGTCTCTCCAGCAT 27
Db      107 CTGCACTCGTGTCTCTCCAGCTT 83

```

```

RESULT 8
US-60-487-610-1136
; Sequence 1136, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CI001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1136
; LENGTH: 904
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-1136

```

```

Query Match          60.0%; Score 19.8; DB 7; Length 904;
Best Local Similarity 72.7%; Pred. No. 29;
Matches 24; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

```

```

Qy      1 CCTGCAAACTCGTGTCTCTCCAGTCAGG 33
Db      782 CCTGCACTCTCCAGTGTCTCCAGATTCTGG 814

```

```

RESULT 9
US-09-908-576-244/C
; Sequence 244, Application US/09908576
; GENERAL INFORMATION:

```

```

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Borstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Geo, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Macher, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tunas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/908,576
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US/09/665,350B
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 244
; LENGTH: 3679
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-908-576-244

```

```

Query Match          60.0%; Score 19.8; DB 5; Length 3679;
Best Local Similarity 84.0%; Pred. No. 37;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy      1 CCTGCAAACTCGTGTCTCTCCAGC 25
Db      1297 CCTGCAAACTGTGTCTCTCCAGC 1273

```

```

RESULT 10
US-60-495-114-16801
; Sequence 16801, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES

```

```

; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16801
; LENGTH: 26000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(26000)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-60-495-114-16801

Query Match          60.0%; Score 19.8; DB 7; Length 26000;
Best Local Similarity 72.7%; Pred. No. 50;
Matches 24; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCTGCAAACTCGTGTCTCTCCAGCATGAGG 33
DB 12108 CCTGCACTCTCCAGTTGTTCCAGCATTTCTGG 12140

RESULT 11
US-60-487-610-19825
; Sequence 19825; Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 19825
; LENGTH: 29102
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-19825

Query Match          60.0%; Score 19.8; DB 7; Length 29102;
Best Local Similarity 72.7%; Pred. No. 51;
Matches 24; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCTGCAAACTCGTGTCTCTCCAGCATGAGG 33
DB 12108 CCTGCACTCTCCAGTTGTTCCAGCATTTCTGG 12140

RESULT 12
US-60-485-450-12335
; Sequence 12335; Application US/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12335
; LENGTH: 29102
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
US-60-485-450-12335

Query Match          60.0%; Score 19.8; DB 7; Length 29102;
Best Local Similarity 72.7%; Pred. No. 51;
Matches 24; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCTGCAAACTCGTGTCTCTCCAGCATGAGG 33
DB 12108 CCTGCACTCTCCAGTTGTTCCAGCATTTCTGG 12140

RESULT 13
US-60-487-610-19824/c
; Sequence 19824; Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 19824
; LENGTH: 60476
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-19824

Query Match          60.0%; Score 19.8; DB 7; Length 60476;
Best Local Similarity 72.7%; Pred. No. 58;
Matches 24; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCTGCAAACTCGTGTCTCTCCAGCATGAGG 33
DB 58060 CCTGCACTCTCCAGTTGTTCCAGCATTTCTGG 58028

RESULT 14
US-60-485-450-12207/c
; Sequence 12207; Application US/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12207
; LENGTH: 60476
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-485-450-12207

Query Match          60.0%; Score 19.8; DB 7; Length 60476;
Best Local Similarity 72.7%; Pred. No. 58;
Matches 24; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCTGCAAACTCGTGTCTCTCCAGCATGAGG 33
DB 58060 CCTGCACTCTCCAGTTGTTCCAGCATTTCTGG 58028

RESULT 15
US-60-495-114-16591/c
```

```

; Sequence 16591, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 16591
; LENGTH: 62312
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(62312)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-60-495-114-16591

```

```

Query Match          60.0%; Score 19.8; DB 7; Length 62312;
Best local Similarity 72.7%; Pred. No. 58;
Matches 24; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

```

```

QY      1  CCTGCAAACTCGTGCTCTCCAGCATGCAAGG 33
          |||||  |||  |||  |||  |||  |||
DB      58061 CCTGCACTCTCCAGTTGTCCAGCATCTGGG 58029

```

```

Search completed: September 16, 2003, 22:48:21
Job time : 59.9505 secs

```

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 17:35:29 ; Search time 394.752 Seconds

(without alignments)
1865.405 Million cell updates/sec

Title: US-09-594-065-2

Perfect score: 18
Sequence: 1 agcagccgcctccttg 18

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pac:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_srs:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_da:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pac:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_srs:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
C 1	18	100.0	798	6 103657	103657 Sequence 5
C 2	18	100.0	1478	14 S74390	S74390 gBpath/UL27
C 3	18	100.0	1488	14 S65444	S65444 UL27-glycop
C 4	18	100.0	1857	14 HS1GBM	K02720 HSV1 (mutan
C 5	18	100.0	2088	6 E03093	E03093 DNA encodin
C 6	18	100.0	2586	6 E00358	E00358 DNA encodin
C 7	18	100.0	2712	6 BD145144	BD145144 Herpes g1
C 8	18	100.0	2712	6 BD165793	BD165793 Vaccine f
C 9	18	100.0	2712	6 BD165795	BD165795 Vaccine f
C 10	18	100.0	2715	6 BD145143	BD145143 Herpes g1
C 11	18	100.0	2715	14 AF097023	AF097023 Human her
C 12	18	100.0	2715	14 AF295528	AF295528 Human her
C 13	18	100.0	2715	14 AF311740	AF311740 Human her
C 14	18	100.0	2715	14 HHU49121	HHU49121 Human herpe
C 15	18	100.0	2809	14 AF258899	AF258899 Human her
C 16	18	100.0	2817	14 BHU14662	BHU14662 Baboon herp
C 17	18	100.0	2818	14 AF021340	AF021340 Human her
C 18	18	100.0	2846	14 HHU12172	HHU12172 Human herpe
C 19	18	100.0	2846	14 HHU12174	HHU12174 Human herpe
C 20	18	100.0	2855	14 HHU12173	HHU12173 Human herpe
C 21	18	100.0	2855	14 HHU12175	HHU12175 Human herpe
C 22	18	100.0	2925	6 E03115	E03115 DNA encodin
C 23	18	100.0	2943	14 H5VGB0	I17281 Sequence 7
C 24	18	100.0	2943	14 H5VGB0	M57388 Simian herp
C 25	18	100.0	2973	14 BHU14663	U14663 Baboon herp
C 26	18	100.0	3098	6 E00883	E00883 Fragment of
C 27	18	100.0	3324	14 HS2GB	M1923 HSV2 glycop
C 28	18	100.0	3461	6 E01195	E01195 DNA sequenc
C 29	18	100.0	3465	6 E03024	E03024 DNA encodin
C 30	18	100.0	3465	6 E03092	E03092 DNA encodin
C 31	18	100.0	3465	6 E03112	E03112 DNA encodin
C 32	18	100.0	3472	6 BD165792	BD165792 Vaccine f
C 33	18	100.0	3472	6 BD165794	BD165794 Vaccine f
C 34	18	100.0	3472	6 108685	I08685 Sequence 1
C 35	18	100.0	3472	6 108686	I08686 Sequence 2
C 36	18	100.0	3472	14 HS2GB3	M15118 HSV-2 (333)
C 37	18	100.0	3640	6 108361	I08361 Sequence 1
C 38	18	100.0	3643	14 HS1GBP	K03541 HSV-1 (Patc
C 39	18	100.0	3715	14 HS2GB2A	M24771 Herpes simp
C 40	18	100.0	3755	6 E03025	E03025 DNA encodin
C 41	18	100.0	3755	6 E03113	E03113 DNA encodin
C 42	18	100.0	3758	14 HS1GB	K01760 HSV1 (KOS)
C 43	18	100.0	3996	6 E03026	E03026 DNA encodin
C 44	18	100.0	3986	6 E03114	E03114 DNA encodin
C 45	18	100.0	3997	6 E00357	E00357 DNA encodin

ALIGNMENTS

RESULT 1

LOCUS 103657

DEFINITION Sequence 5 from Patent US 4642333.

ACCESSION 103657

VERSION 103657.1 GI:268613

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 798)

AUTHORS Person,S.

TITLE Immunologically reactive non-glycosylated amino acid chains of glycoprotein B of herpes virus types 1 and 2

JOURNAL Patent: US 4642333-A 5 10-FEB-1987;

798 bp ss-DNA linear PAT 21-MAY-1993

```

FEATURES
600 Locust La.: State College, PA
Location/Qualifiers
1..798 /organism="unknown"
BASE COUNT 155 a 261 c 254 g 128 t
ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 798;
Best Local Similarity 100.0%; Pred. NO. 3.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AGCAGCGCGCTGTCTTGG 18
|||||
Db 182 AGCAGCGCGCTGTCTTGG 165
|||||

RESULT 2
S74390 1478 bp DNA linear VRL 30-APR-1995
LOCUS S74390
DEFINITION gBpath/UL27=glycoprotein B [herpes simplex virus type 1 HSV-1,
AMgpath, Genomic Mutant, 1478 nt].
S74390
ACCESSION S74390.1 GI:786563
VERSION
KEYWORDS
SOURCE
ORGANISM
Human herpesvirus 1
Viruses; daDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
1 (bases 1 to 1478)
Koetal,M., Bacik,I., Rajcani,J. and Kaerner,H.C.
Replacement of glycoprotein B gene in the herpes simplex virus type
1 strain AMgpath DNA by that originating from nonpathogenic strain
KOS reduces the pathogenicity of recombinant virus
Acta Virol. 38 (2), 77-88 (1994)
95067449
7976866
JOURNAL
MEDLINE
PUBMED
REMARK
Gendank staff at the National Library of Medicine created this
entry [NCBI gBbbag 158378] from the original journal article.
This sequence comes from Fig. 8.
Location/Qualifiers
1..1478
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/db_xref="taxon:10298"
1..1478
1..1478
/partial
/gene="gBpath/UL27"
/note="glycoprotein B"
BASE COUNT 236 a 451 c 528 g 263 t
ORIGIN
Query Match 100.0%; Score 18; DB 14; Length 1478;
Best Local Similarity 100.0%; Pred. NO. 2.9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AGCAGCGCGCTGTCTTGG 18
|||||
Db 348 AGCAGCGCGCTGTCTTGG 365
|||||

RESULT 3
S65444 1488 bp DNA linear VRL 21-APR-2003
LOCUS S65444/c
DEFINITION UL27=glycoprotein B [herpes simplex virus type 1 HSV-1, ANG,
Genomic, 1488 nt].
S65444
S65444.1 GI:415494
VERSION
KEYWORDS
SOURCE
ORGANISM
Human herpesvirus 1
Viruses; daDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
1 (bases 1 to 1488)
REFERENCE
1 (bases 1 to 1488)

```

AUTHORS Yunhas,S.A. and Stevens,J.G.
TITLE Glycoprotein B is a specific determinant of herpes simplex virus type 1 neuroinvasiveness
JOURNAL J. Virol. 67 (10), 5948-5954 (1993)
MEDLINE 93381796
PUBMED 8396662
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbs 137392] from the original journal article.
FEATURES This sequence comes from Fig. 7.
source Location/Qualifiers
1..1488
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/db_xref="taxon:10298"
<1..>1488
/genes="UL27"
<1..>1488
/notes="This sequence comes from Fig. 7"
/codon_start=1
/product="glycoprotein B"
/protein_id="PAP1361.1"
/db_xref="GI:30027723"
/translation="VDFEYARDLTTRKARATPTRNLLTPKFTVANDWPKPSVCT
MTKQVDEMLRSEYSGSPFSSDAISTFTNTLTPELVLDGIDGARDAMR
IPARVATHTVGGQPGVYANGFPIAYGPLNTLAEIYVHEHLEQSRKPNP
PPKASAMASVERIKTSSIEPRADLTQYTHIQRVNDMGRVLAACEQNHETLM
NEARKLVNDAIASATVGRVRSARLGVMAVSTCPVAADNVAVONSRLSRGACT
SRPLVSFRYEDQGLVEYQGENNELRLTDAIEPCVGRHRYFTFGGVYFEYVAV
SHQSRADITTVSTFDLNTLMLEDHEFVPLEYTRHEIKDSGLIDYEVORNLHD
LRFADITVIAHDANAMPAFGALGFEFGMDLRAVGVKMGIVGVSAVSGVSPFM
SNPFGALAVGLVLIAGLAAPFAPRYVMRLQSNPKALYPLTTEL"
BASE COUNT 260 a 530 c 459 g 239 t
ORIGIN
Query Match 100.0%; Score 18; DB 14; Length 1488;
Best Local Similarity 100.0%; Pred. No. 2,9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1136 AGCAGCGCGCTGCTCTTG 1119
Qy 1 AGCAGCGCGCTGCTCTTG 18
|||||
|||||
RESULT 4
HSIGEM/c HSIGEM 1857 bp DNA linear VRL 02-AUG-1993
LOCUS HSV1 (mutant strain ts85), glycoprotein B (gb) gene.
DEFINITION K02720
ACCESSION K02720.1 GI:330087
KEYWORDS glycoprotein.
REMARKS Human herpesvirus 1
Human herpesvirus 1
Virus(es): dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
1 (bases 1 to 1857)
Bzik,D.J., Fox,B.A., Deluca,N.A. and Pearson,S.
Nucleotide sequence of a region of the herpes simplex virus type 1
gb glycoprotein gene: mutations affecting rate of virus entry and
cell fusion
REFERENCE Virology 137 (1), 185-190 (1984)
AUTHORS 84301857
TITLE 6089415
JOURNAL Original source text: HSV1 strain ts85 DNA, clone pTBG-B51.
MEDLINE HSV1 strain ts85 is a mutant of strain HPEM that is temperature
PUBMED dependent for virus growth, production of gb, and fusion of Vero
COMMENT and HEL cells. The syn and ts lesions of ts85 are separable by
recombination. The syn defect being located between 0.345 and 0.355
map units, and the ts defect being located between 0.360 and 0.368
map units. The rate of entry determinant of ts85 is genetically
separable from both the syn and ts lesions and is located between
0.350 and 0.360 map units, which is entirely within the
glycoprotein B gene (0.348 to 0.366 map units).

Strain ts85 has a total of 12 base substitutions within the coding region for gB when compared to strain KOS. Of these, only two cause amino acid substitutions; a 't' to 'c' transition at bp 695 causes a Val to Ala substitution, and a 'g' to 'a' transition at bp 1610 causes an Arg to His substitution. Because the rate-of-entry determinant has been mapped to between 0.350 and 0.360 map units, the transition at bp 695 must be responsible for the difference in rate-of-entry between ts85 and KOS.

FEATURES

source

1..1857
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/db_xref="taxon:10298"
<1..1752
/note="glycoprotein B"
/codon_start=1
/protein_id="AAA45777.1"
/db_xref="GI:330088"

CDS

Query Match 100.0%; Score 18; DB 14; Length 1857;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

BASE COUNT 358 a 624 c 576 g 299 t
ORIGIN 1 bp upstream of SalI site; map coordinate 0.360.

Query Match 100.0%; Score 18; DB 14; Length 1857;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
E03093 2088 bp DNA linear PAT 29-SEP-1997
LOCUS E03093
DEFINITION DNA encoding truncated gB of herpes simplex virus type1 (HSV-1).
ACCESSION E03093
VERSION E03093.1 GI:2171311
KEYWORDS JP 1991218397-A/2.
SOURCE Human herpesvirus 1
ORGANISM Human herpesvirus 1
Virusess; dadNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
1 (bases 1 to 2088)
Fujisawa,Y., Hinuma,K., Asakawa,N. and Otake,S.
POLYPEPTIDE AND ITS PRODUCTION
Patent: JP 1991218397-A 2 25-SEP-1991;
TAKEDA CHEM IND LTD
OS Herpes simplex virus type1 (HSV-1)
PN JP 1991218397-A/2
PD 25-SEP-1991
PF 21-JUN-1990 JP 1990161448
PR 22-JUN-1989 JP 89P 158238, 30-NOV-1989 JP 89P 308941 PI
FUJISAWA YUKIO, HINUMA KUNIKI, ASAKAWA NAOKI, OTAKA SACHIKO PC
C07K13/00, C12N1/19, C12N15/38, C12P21/02//A61K39/245, (C12N1/19, PC
C12R1:865),
PC (C12P21/02, C12R1:865);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=MiYama;
CC *source: clone=PHS106 delta Tth;
FH Key Location/Qualifiers

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FT CDS 1..2085
FT /product='truncated gB of HSV-1' FT 3'UTR
FT 2086..2088
FT mat_peptide 1..2082
FT /product='truncated gB of HSV-1'.
Location/Qualifiers
1..2088
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/db_xref="taxon:10298"

BASE COUNT 417 a 751 c 620 g 300 t
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 2088;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 18; DB 6; Length 2088;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
E00358 2586 bp DNA linear PAT 29-SEP-1997
LOCUS E00358
DEFINITION DNA encoding glycoprotein B of herpes simplex virus type 2.
ACCESSION E00358
VERSION E00358.1 GI:2168645
KEYWORDS JP 1985115529-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2586)
AUTHORS Sutanee,P.
TITLE NON-GLYCOSYLATED AMINO ACID CHAIN IMMUNOLOGICALLY REACTIVE TO GLYCOPROTEIN OF HERPES VIRUS 1 AND 2
JOURNAL Patent: JP 1985115529-A 2 22-JUN-1985;
SUTANEE PAASON
COMMENT OS herpes simplex virus type 2
PN JP 1985115529-A/2
PD 22-JUN-1985
PF 23-JUN-1984 JP 1984129915
PR 23-JUN-1983 US 83 506986, 16-SEP-1983 US 83 532996 PI
SUTANEE PAASON
PC A61K39/245, C07H21/04, C07K15/14, C12N15/00//C12P21/02; CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HG 52;
CC *source: clone=p52BX;
CC *source: map_position=0.386-0.348;
FH Key Location/Qualifiers
FT CDS 1..2727
FT /product='glycoprotein B of herpes simplex FT
FT virus type 2'.
Location/Qualifiers
1..2586
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES
source
1..2586
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 481 a 886 c 841 g 378 t
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 2586;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 18; DB 6; Length 2586;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 7
BD145144/c
LOCUS BD145144 2712 bp DNA linear PAT 17-JAN-2003
DEFINITION Herpes simplex virus VP16 vaccine.
ACCESSION BD145144
VERSION BD145144.1 GI:27850902
KEYWORDS JP 2002136297-A/4.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2712)
AUTHORS Burke, R.L. and Sekulovich, R.E.
TITLE Herpes simplex virus VP16 vaccine
JOURNAL Patent: JP 2002136297-A 4 14-MAY-2002;
CHIRON CORP
COMMENT OS HSV gB1
PN JP 2002136297-A/4
PD 14-MAY-2002
PF 10-SEP-2001 JP 2001274335
PR 02-AUG-1990 US 561,528
PI RAE LYN BURKE, ROSE E SEKULOVICH
PC C12N15/09, A61K39/245, A61P31/12, C07K14/03, C12N15/00,
A61K37/02
CC Herpes simplex virus VP16 vaccine
FH Key Location/Qualifiers
FT CDS (1)..(2712).
FEATURES
source 1..2712
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 513 a 949 c 853 g 397 t
ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 2712;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGCAGGCCGCTGCTCTTG 18
|||||
DB 2096 AGCAGGCCGCTGCTCTTG 2079
|||||
RESULT 8
BD165793/c
LOCUS BD165793 2712 bp DNA linear PAT 17-JAN-2003
DEFINITION Vaccine for use in the therapeutic treatment of HSV.
ACCESSION BD165793
VERSION BD165793.1 GI:27871605
KEYWORDS JP 2002167398-A/2.
SOURCE Herpes simplex virus 7
ORGANISM Herpes simplex virus 7
REFERENCE 1 (bases 1 to 2712)
AUTHORS Burke, R.L., Pachl, C. and Valenzuela, P.D.T.
TITLE Vaccine for use in the therapeutic treatment of HSV
JOURNAL Patent: JP 2002167398-A 2 11-JUN-2002;
CHIRON CORP
COMMENT OS Herpes simplex virus 7
PN JP 2002167398-A/2
PD 11-JUN-2002
PF 12-SEP-2001 JP 2001277340
PR 20-OCT-1986 US 921,213
PI RAE LYN BURKE, CAROL PACHL, PABLO D T VALENZUELA PC
C07K14/035, A61K39/245, A61P31/22, C12N15/09, C12N15/00, C12N15/00
CC Vaccine for use in the therapeutic treatment of HSV FH Key
Location/Qualifiers
FT CDS (1)..(2712).
FEATURES
source Location/Qualifiers
FT CDS (1)..(2712).
FEATURES

source 1..2712
/organism="Herpes simplex virus 7"
/mol_type="genomic DNA"
/db_xref="taxon:57277"
BASE COUNT 490 a 938 c 884 g 400 t
ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 2712;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGCAGGCCGCTGCTCTTG 18
|||||
DB 2090 AGCAGGCCGCTGCTCTTG 2073
|||||
RESULT 9
BD165795/c
LOCUS BD165795 2712 bp DNA linear PAT 17-JAN-2003
DEFINITION Vaccine for use in the therapeutic treatment of HSV.
ACCESSION BD165795
VERSION BD165795.1 GI:27871607
KEYWORDS JP 2002167398-A/4.
SOURCE Herpes simplex virus 7
ORGANISM Herpes simplex virus 7
REFERENCE 1 (bases 1 to 2712)
AUTHORS Burke, R.L., Pachl, C. and Valenzuela, P.D.T.
TITLE Vaccine for use in the therapeutic treatment of HSV
JOURNAL Patent: JP 2002167398-A 4 11-JUN-2002;
CHIRON CORP
COMMENT OS Herpes simplex virus 7
PN JP 2002167398-A/4
PD 11-JUN-2002
PF 12-SEP-2001 JP 2001277340
PR 20-OCT-1986 US 921,213
PI RAE LYN BURKE, CAROL PACHL, PABLO D T VALENZUELA PC
C07K14/035, A61K39/245, A61P31/22, C12N15/09, C12N15/00, C12N15/00
CC Vaccine for use in the therapeutic treatment of HSV FH Key
Location/Qualifiers
FT CDS (1)..(2712).
FEATURES
source 1..2712
/organism="Herpes simplex virus 7"
/mol_type="genomic DNA"
/db_xref="taxon:57277"
BASE COUNT 514 a 949 c 853 g 396 t
ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 2712;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGCAGGCCGCTGCTCTTG 18
|||||
DB 2099 AGCAGGCCGCTGCTCTTG 2082
|||||
RESULT 10
BD145143/c
LOCUS BD145143 2715 bp DNA linear PAT 17-JAN-2003
DEFINITION Herpes simplex virus VP16 vaccine.
ACCESSION BD145143
VERSION BD145143.1 GI:27850901
KEYWORDS JP 2002136297-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2715)
AUTHORS Burke, R.L. and Sekulovich, R.E.
TITLE Herpes simplex virus VP16 vaccine

JOURNAL Patent: JP 2002136297-A 3 14-MAY-2002;
CHIRON CORP
COMMENT OS HSV GB2
PN JP 2002136297-A/3
PD 14-MAY-2002
PE 10-SEP-2001 JP 2001274335
PR 02-AUG-1990 US 561,528
PI KAE LYNN BURKE, ROSE B SEKULOVICH
PC C12N15/09, A61K39/00, A61K39/245, A61P31/12, C07K14/03, C12N15/00,
PC A61K37/02
CC Herpes simplex virus VP16 vaccine
FT Key Location/Qualifiers
CDS (1)..(2715).

BASE COUNT 491 a 938 c 886 g 400 t
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 2715;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCAGCGCGCTGCTCTTG 18
Db 2090 AGCAGCGCGCTGCTCTTG 2073

RESULT 11
AF097023/c 2715 bp DNA linear VRL 07-APR-2000
LOCUS Human herpesvirus 1 strain HSZP glycoprotein B (UL27) gene,
DEFINITION complete cds.
ACCESSION AF097023.1 GI:6165611
VERSION AF097023.1
KEYWORDS Human herpesvirus 1
SOURCE Human herpesvirus 1
ORGANISM Viruses; dAdNA viruses, no RNA stage; Herpesviridae;
REFERENCE 1 (bases 1 to 2715)
AUTHORS Kosovsky, J., Vojvodova, A., Oravcova, I., Kudelova, M., Matls, J. and Rajcani, J.
TITLE Herpes simplex virus 1 (HSV-1) strain HSZP glycoprotein B gene: comparison of mutations among strains differing in virulence
JOURNAL Virus Genes 20 (1), 27-33 (2000)
PUBMED 20227371
REFERENCE 2 (bases 1 to 2715)
AUTHORS Kosovsky, J., Vojvodova, A., Oravcova, I., Majorova, J., Kudelova, M. and Rajcani, J.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-1998) Institute of Virology, Slovak Academy of Sciences, Dubravska cesta 9, Bratislava 842 46, Slovakia
FEATURES
source
1..2715
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/strain="HSZP"
/db_xref="taxon:10298"
1..2715
/gene="UL27"
1..2715
/gene="UL27"
/function="essential for fusion between virion envelope and cell membrane"
/note="envelope glycoprotein, gB"
/codon_start=1
/product="glycoprotein B"
/protein_id="AA04615.1"
/db_xref="GI:6165612"

/translation="NRQAPARGCGFVVMALLGLTLGVASAPSSPGTGVAAAT
QAANGPAPAPALGAPGTDGPKKKKKKKKKPPPPAGDNATVAAGATLREHLR
DIKASTANAFVCPPTGATVQFQEPKRCPTPEGGVYTGIAVFEENLAPYFK
ATMYKDVTSQVWFGRYSQPMGI FEEBAPVPFESEVLDIKNAKVCSTAYVNNL
ETTAPHRDHTDMELKXPNATRTSRCHMTDLKYNBSRVARFRCTACTYCEBV
DARSYTPDEFLATGDFVYNSPFYGEHSTHTETSTIAADFKQDGFATDLTTKA
RATAPPTRNLLTPKFTVAMDVPFRPSVCTWKQOEDMLRSEYGSFRSSDAIS
TPTTNLLEYPLSRVLDGCIQKARDANDRLEFARVYNAHAKVOPOYYLANGFLI
AYOPLSTNLAEVLEHLEBQSRKPPPTPPGASANAVERIKTSSIEFARLOF
TYNHIQRHVNMDLGRVAIALMCELQKHLETLNBEAKLNALASATVGRVSARMGD
VMAVSTCVPAADNYIVONSMGSRPACVSRPLVSPRYEGRPLVBCGLDENNELR
LTRDAIBECTVGRHRYFTFGGCTVYFEETAYSHQLSRADITTVTSFTIDNTLMEDHE
FVPLEVYTRHEIKDSGLDYTEVQRNOLDLRFADIDTVIHADNANAPALGAEFE
GMGDIGRAVGVKVMGIVGVSAVGSFMSNPFGALAVGLVLAGIAAFAFRVY
KLGQSNPKMALPVLTKELKPTNPDSAGEGCEGDFDEAKLAEAREMIRYALVVSAM
EHTEHAKKKGTSLLSAKVYDMVRKRNNTYTVQVPMKDGADBDL"

BASE COUNT 516 a 943 c 856 g 400 t
ORIGIN

Query Match 100.0%; Score 18; DB 14; Length 2715;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCAGCGCGCTGCTCTTG 18
Db 2099 AGCAGCGCGCTGCTCTTG 2082

RESULT 12
AF295528/c 2715 bp mRNA linear VRL 01-SEP-2001
LOCUS Human herpesvirus 2 glycoprotein B2 (GB2) mRNA, complete cds.
DEFINITION AF295528
ACCESSION AF295528.1 GI:15428575
VERSION AF295528.1
KEYWORDS Human herpesvirus 2
SOURCE Human herpesvirus 2
ORGANISM Viruses; dAdNA viruses, no RNA stage; Herpesviridae;
REFERENCE 1 (bases 1 to 2715)
AUTHORS Lee, H.H.
TITLE Sequencing and high level expression of the HSV-2 strain G glycoprotein B gene using a baculovirus HcNPV
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2715)
TITLE Lee, H.H.
JOURNAL Direct Submission
JOURNAL Submitted (12-AUG-2000) Department of Biology, Konkuk University, 1 Hwayang-dong, Kwangjin-gu, Seoul 143-701, Korea
FEATURES
source
1..2715
/organism="Human herpesvirus 2"
/mol_type="mRNA"
/strain="G"
/db_xref="taxon:10310"
1..2715
/gene="gB2"
1..2715
/gene="gB2"
/codon_start=1
/product="glycoprotein B2"
/protein_id="AAK97852.1"
/db_xref="GI:15428575"
/translation="NRGGGLICALVGLVLAASAAAPAPAPASGVAATVAANG
GPASRPVPSPTATKARKKTKKPEPEPEPPPDANAATVAAGATLRAHLREIVE
NADAPFVCPPTGATVQFQEPKRCPTPEGGVYTGIAVFEENLAPYFKATMY
KDVYTSQVWFGRYSQPMGI FEEBAPVPFESEVLDIKNAKVCSTAYVNNL
HDDHETDMELKPAKVAATRTSRCHMTDLKYNBSRVARFRCTACTYCEBVARSV
YPCBEVATGDDVYNSPFYGEHSTHTETSTIAADFKQDGFATDLTTKARATSP
TTRNLTTPKFTVAMDVPFRPSVCTWKQOEDMLRSEYGSFRSSDAISSTPTT
NLTOYLSRVDGDTIGKARBEADIMRFRKRNATHIRVGOQYTLATGGFLIATOP
LSNTLAELVREYMEQDRKRNATPAPLRAPSPANASVERIKTSSIEFARLOFTYN
HIOHVNMDLGRVAIALMCELQKHLETLNBEAKLNALASATVGRVSARMGDVMA

BASE COUNT 492 a 939 c 866 g 358 t
ORIGIN
VSTCTGPAAPADNRIVYVONSMRVSRSRPGTCSRSRLVSFRREDGGPFLIEQDGGENNELTLTR
DALEPCTVGHRRYFI FGGGYYFEEVAYSHOLSRADYTTSTFDILGNTLEEDH IYPT
LEVYTRHRIKDSGLLDYTEVORRNQHLDRADYDIT IADANVAMVAGCA FPEGGQD
DLGTAUVKVMGQGVAVSAVGSVSPMSNFGVAGLAVLAVLAVMAFAFRVGT
QRNMFKMLPLLTKELELSTDPGCVGGEBCAGCGDEALAEARMIYMYLVASAM
ERTHKARKKKTSTALLSSTKVTNMYLRKKNKRVPLHNEDEAGDEDEL"

	Query Match	Score 18;	DB 14;	length 2715;
Best Local Similarity	100.0%	100.0%	Pred. No. 2,7e+02;	
Matches 18;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1 AGCAGCGCGCTGTCTCTTG 18			
Db	2090 AGCAGCGCGCTGTCTCTTG 2073			

[illegible]

REFERENCE	1 (bases 1 to 2715)
AUTHORS	Petel, P. E., Fildberg, A., Parish, M. L. and Spear, P. G.
TITLE	Cell fusion induced by herpes simplex glycoproteins gB, gD, and gH/gL requires a gD receptor but not necessarily heparan sulfate
JOURNAL	Virology 279 (1), 313-324 (2001)
MEDLINE	21066717
PUBMED	11145912
REFERENCE	2 (bases 1 to 2715)
AUTHORS	Petel, P. E., Fildberg, A., Parish, M. L. and Spear, P. G.
TITLE	Direct Submision
JOURNAL	Submitted (05-OCT-2000) Microbiology-Immunology, Northwestern University Medical School, 320 E. Superior Ave., Ward 6-241, Chicago, IL 60611, USA

```

FEATURES
SOURCE
LOCATION/Qualifiers
1. .2715
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/strain="KOS"
/db_xref="taxon:10298"
1. .2715
CDS

```

BASE COUNT	521 a	948 c	849 g	397 e
------------	-------	-------	-------	-------

Query Match	100.0%	Score 18	DB 14	Length 2715
Best Local Similarity	100.0%	Pred. No. 2.7e+02		
Matches 18	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	AGCAGCGCGCTGCTGCTG 18		
DB	2099	AGCAGCGCGCTGCTGCTGCTG 2082		

RESULT 14					
HHU49121/c	HHU49121	2715 bp	DNA	linear	VRL 08-MAR-1996
LOCUS	Human herpesvirus 1 strain ANG path.				
DEFINITION	complete cds.				
ACCESSION	U49121				
VERSION	U49121.1	GI:1218049			

ORGANISM	REFERENCE	TITLE	AUTHORS	JOURNAL	FEATURES
Human herpesvirus 1 Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Simplexvirus.	1 (bases 1 to 2715)		Holland, T.C. and Sabathin-Langrood, A.		
	Direct Submission				
	Submitted (13-FEB-1996)				
	State University, 540 E. Canfield Ave., Detroit, MI 48201, USA				
	location/Qualifiers				

```

gene
    /mol_type="genomic DNA"
    /strain="ANG path"
    /db_xref="taxon:10296"
    .2715
    /gene="gb"
    /note="HSV-1 UL27 open reading frame"
    1..2715
CDS

```

[illegible]

RESULT 15

DEFINITION Human herpesvirus 1 glycoprotein B (UL27) gene, complete cds.
ACCESSION AF259899
VERSION AF259899.1 GI:7839508
KEYWORDS
SOURCE
ORGANISM Human herpesvirus 1
Human herpesvirus 1
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
REFERENCE
AUTHORS 1 (bases 1 to 2809)
TITLE Ling, J.-Y., Chen, T.-M. and Stroop, W.G.
A viral genetic element involved in a rabbit model of herpes
simplex virus-induced epileptiform seizures
JOURNAL Unpublished
REFERENCES 2 (bases 1 to 2809)
AUTHORS Ling, J.-Y., Chen, T.-M. and Stroop, W.G.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2000) Ophthalmology, Mail Slot 523, University of
Arkansas for Medical Sciences, 4301 West Markham, Little Rock, AR
72205, USA

FEATURES
source
1..2809
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/strain="+GC"
/db_xref="taxon:10298"
38..2752
/gene="UL27"
38..2752
/gene="UL27"
/codon_start=1
/product="glycoprotein B"
/protein_id="AAF70301.1"
/db_xref="GI:7839509"
/translation="MRQCAPARGCMFVVMALIGLTIGLVNAAAPSSRPTGVAAT
QAANGPATRPAPAPGATGPTGKAKKKKPPPPPCGAGNATVAAGHATLREHL
DIAKANTANFVCPPTGATVQGEQPRCPTRREGQNTYTGIAVFERENTAPYKFK
ATMYKQDVTSQWFGHRSQPMGI.FEDRAPVPFBEVIDKINAKGCRSTAKYVNNL
ETTAHRDHDHETDMLKPPANAAATRTSRGHTTDLKYNPSRVAFAFRGTTVNCIVEV
DARSVPYNEFVATGDFYMSPFYGRSGSHETHTSYAADRFKQVDGFYADLTTKA
RATAPTRNLTTPKFTVAMDVPPRPSPVCMTKQVDEMTLSREYGSFRSSDAIS
TPTFTNLTEYPLSRVDLSDCTGKARDAMDRI.FARRYNATHIKVGQPOYTLANGFLLI
AYQPLSNLTAEIYVREHLREQSRKPPNTPPPGASANASVERIKTTSIEFARLQF
TYNNHQHVNDMLGRVAIAWELQNHETLMEARKLNPNALASVTVGRVRSARMIGD
VMAVSTCVPVADNIVIONSMRISSRPGACYSRPLSPRYEDQPLVEGQLENNEILR
LTRDAIECTVGRHRYFTFGGGYVFEYAVSHOLSRADITTVSTFIDNTIMLEDEH
FVPLEVYRPHETKDSGLDYTEVQRNQLHDLRFADIDVTIHADANAAAFAGLGAFFE
GMGDLGRAVGVKVMGIVGVSVASGVSSFMSPFGALVGLVLAGLAAPFARFV
MRLQSNPMKALPLITTKELKNTPNPDASGESEGGDFEAKLAEAREMIRYALVSAM
EHTENKAKKKGTSALISAKVTDMVRKRNTNYTQVPMKSDADEDDL"

BASE COUNT 536 a 974 c 883 g 416 t
ORIGIN

Query Match 100.0%; Score 18; DB 14; Length 2809;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 AGCAGGCCGCTGCTTG 18
|||||
2136 AGCAGGCCGCTGCTTG 2119

Search completed: September 16, 2003, 19:14:13
Job time : 395.752 secs

BEST AVAILABLE COPY

DR WPI; 1991-328397/45.
 DR P-PSDB; AAR14666.
 XX HSVGB polypeptide(s) obcd. by recombinant DNA techniques -
 PT useful as vaccines against HSV and in diagnosis, can be produced
 PT cheaply and safely.
 PS Claim 3; Fig 7; 24pp; Japanese.
 CC The sequence encodes a truncated form of the HSVGB polypeptide.
 CC The recombinant protein can be used to prepare vaccines for
 CC prophylaxis of HSV infection and for use in diagnostic kits.
 CC See also AAQ14478.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 SQ Sequence 2088 BP; 417 A; 751 C; 620 G; 300 T; 0 other;
 Query Match 100.0%; Score 18; DB 12; Length 2088;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 AGCAGCGCGCTGCTCTTG 18
 DB 2012 AGCAGCGCGCTGCTCTTG 1995
 RESULT 2
 ID AAV62158 standard; DNA; 2645 BP.
 AC AAV62158;
 DT 18-DEC-1998 (first entry)
 XX
 DE HSV-2 strain SB5 Contig ID 14 DNA sequence.
 XX
 KW HSV-2 strain SB5; immunological response induction; therapy;
 KW antiviral identification; viral protein inhibitor; ss.
 XX
 OS Herpes simplex virus type 2.
 XX
 FH Key Location/Qualifiers
 FT CDS complement (81..2645)
 FT /*tag= a
 FT /product= "ORF1 protein"
 XX
 PN W09820016-A1.
 PD 14-MAY-1998.
 XX
 PF 31-OCT-1997; 97WO-US20016.
 XX
 PR 09-JUN-1997; 97US-0048018.
 PR 04-NOV-1996; 96US-0030279.
 XX
 PA (SMITK) SMITHKLINE BEECHAM CORP.
 XX
 PI Chan JY, Dabrowski-Amaral CB, Delvecchio AM, Dillon SB;
 PI Esser KM, Leary JY;
 XX
 DR WPI; 1998-286847/25.
 DR P-PSDB; AAW72113.
 PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention
 PT and treatment of infection or inducing immunological response in
 PT mammal
 XX
 PS Claim 1; Page 406-407; 748pp; English.
 CC This sequence represents a Herpes simplex virus type-2 (HSV-2) DNA
 CC sequence of the invention. This sequence was isolated from HSV-2 strain
 CC SB5 (deposited as ATCC VR-2546), is designated Contig ID 14, and encodes
 CC a HSV-2 protein. The protein can be used for the treatment or

CC prevention of disease, to induce an immunological response in a mammal or
 CC to identify inhibitors, activators or novel antivirals. Antagonists of
 CC the proteins can be used to inhibit a viral polypeptide. The DNA sequence
 CC or a vector containing it can also be used to induce an immunological
 CC response in a mammal.
 XX
 SQ Sequence 2645 BP; 404 A; 837 C; 892 G; 512 T; 0 other;
 Query Match 100.0%; Score 18; DB 19; Length 2645;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 AGCAGCGCGCTGCTCTTG 18
 DB 706 AGCAGCGCGCTGCTCTTG 723
 RESULT 3
 ID AAN71303/C
 ID AAN71303 standard; DNA; 2712 BP.
 AC AAN71303;
 DT 25-MAR-2003 (updated)
 DT 30-APR-1991 (first entry)
 XX
 DE Herpes Simplex Virus-1 GB from pKBXX.
 XX
 KW Vaccine; prophylaxis; treatment; Herpes Simplex Virus-1;
 KW glycoprotein; gb; ss.
 XX
 OS Herpes simplex virus type 1 (HOS).
 XX
 PN US4642333-A.
 PD 10-FEB-1987.
 XX
 PF 20-JUN-1984; 84US-0622496.
 XX
 PR 20-JUN-1984; 84US-0622496.
 XX
 PA (PERS/) PERSON S.
 XX
 PI Person S;
 XX
 DR WPI; 1987-056354/08.
 DR P-PSDB; AAN71135.
 PT Amino acid chain of glycoprotein B of HSV-1 and 2 - prep'd. as
 PT recombinant and used for vaccines for herpes simplex virus types 1
 PT and 2.
 PS Example; Table 2; 16pp; English.
 XX
 CC A pure non-glycosylated amino acid (AA) chain comprising a sequence
 CC corresponding to that occurring in glycoprotein B of HSV-1 or HSV-2
 CC virus which is antigenic to HSV-1 of HSV-2, which contains no more
 CC than 750 AA residues, and which includes AA residues 135-649
 CC inclusive is claimed. It can be used to produce vaccines for
 CC prophylaxis and treatment of HSV-1 and HSV-2.
 CC See AAN71359 for the HSV-2 sequence.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC
 SQ Sequence 2712 BP; 519 A; 945 C; 849 G; 397 T; 2 other;
 Query Match 100.0%; Score 18; DB 8; Length 2712;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 AGCAGCGCGCTGCTCTTG 18
 DB 2096 AGCAGCGCGCTGCTCTTG 2079

RESULT 4
 AAN71399/c
 ID AAN71399 standard; DNA; 2724 BP.
 AC AAN71399;
 XX
 XX
 DT 25-MAR-2003 (updated)
 DT 30-APR-1991 (first entry)
 XX
 XX
 DE Herpes Simplex Virus-2 gB from p52BXX.
 XX
 KM Vaccine; prophylaxis; treatment; Herpes Simplex Virus-1;
 KM glycoprotein; gB; ss.
 XX
 OS Herpes simplex virus type 2 HGS2.
 XX
 PN US642333-A.
 XX
 PD 10-FEB-1987.
 XX
 PF 20-JUN-1984; 84US-0622496.
 XX
 PR 20-JUN-1984; 84US-0622496.
 XX
 PA (PERS/) PERSON S.
 XX
 PI Person S;
 XX
 DR WPI; 1987-056354/08.
 DR P-PSDB; AAP71136.
 XX
 PT Amino acid chain of glycoprotein B of HSV-1 and 2 - prepd. as
 PT recombinant and used for vaccines for herpes simplex virus types 1
 PT and 2.
 XX
 PS Example; Table 2; 16pp; English.
 XX
 CC 'N' indicated nucleotide to be determined (corresp. to residues
 CC 17-41 of HSV-1 gB).
 CC A pure non-glycosylated amino acid (AA) chain comprising a sequence
 CC corresponding to that occurring in glycoprotein B of HSV-1 or HSV-2
 CC virus which is antigenic to HSV-1 of HSV-2, which contains no more
 CC than 750 AA residues, and which includes AA residues 135-649
 CC inclusive is claimed. It can be used to produce vaccines for
 CC prophylaxis and treatment of HSV-1 and HSV-2.
 CC See AAN71303 for the HSV-1 sequence.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC
 XX
 SQ Sequence 2724 BP; 494 A; 911 C; 854 G; 390 T; 75 other;
 Query Match 100.0%; Score 18; DB 8; Length 2724;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGCAGCGCGCTGCTCTTG 18
 Db 2099 AGCAGCGCGCTGCTCTTG 2082
 RESULT 5
 AAN50517/c
 ID AAN50517 standard; DNA; 2727 BP.
 AC AAN50517;
 XX
 XX
 DT 21-NOV-1991 (first entry)
 XX
 DE Sequence encoding the glycoprotein B (gB) of Herpes simplex virus
 DE type 2 (HSV-2).
 XX
 KM Vaccine; glycoprotein B; ss.
 XX

OS Herpes simplex virus type 2.
 XX
 XX Key Location/Qualifiers
 FT CDS 1..2727
 FT /*tag= a
 XX
 XX EP133063-A.
 XX
 PD 13-FEB-1985.
 XX
 PE 22-JUN-1984; 84EP-0401312.
 XX
 PR 16-SEP-1983; 83US-0532996.
 PR 23-JUN-1983; 83US-0506986.
 PR 20-JUN-1984; 84US-0622496.
 XX
 PA (PERS/) PERSON S.
 XX
 PI Person S;
 XX
 DR WPI; 1985-039636/07.
 XX
 PT New amino acid chains of glyco-protein B of herpes simplex virus
 PT - prepd. by recombination DNA methods for use in vaccines
 XX
 PS Disclosure; Table 2, Page 35-35C; 40pp; English.
 XX
 CC The inventors claim a non-glycosylated amino acid chain comprising a
 CC sequence corresponding to that occurring in glycoprotein B of HSV-1
 CC or HSV-2. Preferably, when the chain comprises a sequence corresp.
 CC to that occurring in glycoprotein B of HSV-2, it contains 4-750 AA
 CC residues. The chain esp. has a molecular wt. of 65000 daltons and
 CC includes AA residues 165-629 from the HSV-2 glycoprotein B sequence.
 XX
 SQ Sequence 2727 BP; 493 A; 913 C; 855 G; 391 T; 75 other;
 Query Match 100.0%; Score 18; DB 6; Length 2727;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGCAGCGCGCTGCTCTTG 18
 Db 2102 AGCAGCGCGCTGCTCTTG 2085
 RESULT 6
 AAV62146/c
 ID AAV62146 standard; DNA; 2841 BP.
 AC AAV62146;
 XX
 DT 07-DEC-1998 (first entry)
 XX
 DE HSV-2 strain SB5 Contig ID 91 DNA sequence.
 XX
 KM HSV-2 strain SB5; immunological response induction; therapy;
 KM antiviral identification; viral protein inhibitor; ss.
 XX
 OS Herpes simplex virus type 2.
 XX
 XX Key Location/Qualifiers
 FT CDS 358..2745
 FT /*tag= a
 XX
 XX MO9820016-A1.
 XX
 PD 14-MAY-1998.
 XX
 PE 31-OCT-1997; 97WO-US20016.
 XX
 PR 09-JUN-1997; 97US-0049018.
 PR 04-NOV-1996; 96US-0030279.
 XX

PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB;
 PI Esser KM, Leary JJ;
 XX WPI; 1998-286847/25.
 DR P-PSDB; AAW72062.
 XX
 PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention
 PT and treatment of infection or inducing immunological response in
 PT mammal
 XX
 PS Claim 1; Page 284-285; 748pp; English.
 XX
 CC This sequence represents a Herpes simplex virus type-2 (HSV-2) DNA
 CC sequence of the invention. This sequence was isolated from HSV-2 strain
 CC SBS (deposited as ATCC VR-2546), is designated Contig ID 91, and encodes
 CC a HSV-2 protein. The protein can be used for the treatment or
 CC prevention of disease, to induce an immunological response in a mammal or
 CC to identify inhibitors, activators or novel antivirals. Antagonists of
 CC the protein can be used to inhibit a viral polypeptide. The DNA sequence
 CC or a vector containing it can also be used to induce an immunological
 CC response in a mammal.
 CC
 SQ Sequence 2841 BP; 577 A; 921 C; 915 G; 425 T; 3 other;
 Query Match 100.0%; Score 18; DB 19; Length 2841;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AGCAGCGCGCTGCTCTTG 18
 DB 2120 AGCAGCGCGCTGCTCTTG 2103
 RESULT 7
 AA014455/C
 ID AA014455 standard; DNA; 2925 BP.
 XX
 AC AA014455;
 XX
 DT 25-MAR-2003 (updated)
 DT 21-JAN-1992 (first entry)
 XX
 DE HSV surface antigens gB/gD fusion protein gene.
 XX
 KM Herpes simplex virus; vaccine; ss.
 XX
 OS Herpes simplex virus.
 XX
 PI Key Location/Qualifiers
 PT CDS 1..2925
 FT /*tag= a
 XX
 PN JP03220200-A.
 XX
 PD 27-SEP-1991.
 XX
 XX 29-NOV-1990; 90JP-0325474.
 PP
 XX 30-NOV-1989; 89JP-0308942.
 PR
 XX 29-NOV-1990; 90JP-0325474.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 DR WPI; 1991-329235/45.
 DR P-PSDB; AAR14680.
 XX
 PT Polypeptide for herpes simplex virus vaccine prodr. - obd. by
 PT culturing transformant of recombinant DNA coding fused protein of
 PT virus surface antigen
 XX
 PS Claim 6; Fig 8; 16pp; Japanese.

XX
 CC The DNA sequence encodes a fusion protein comprising herpes simplex
 CC virus (HSV) surface antigens gB and gD. This fusion protein shows
 CC the same physiological activity as surface antigen prep. from HSV
 CC infected cells. It can be used in a vaccine for the prevention of
 CC HSV infection.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 2925 BP; 579 A; 1061 C; 848 G; 437 T; 0 other;
 Query Match 100.0%; Score 18; DB 12; Length 2925;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AGCAGCGCGCTGCTCTTG 18
 DB 2012 AGCAGCGCGCTGCTCTTG 1995
 RESULT 8
 AAT16480/C
 ID AAT16480 standard; DNA; 2943 BP.
 XX
 AC AAT16480;
 XX
 DT 11-MAY-1996 (first entry)
 DT
 XX
 DE SAb virus gB glycoprotein coding sequence.
 XX
 KM Herpes simian monkey SAb virus gB glycoprotein; immunoassay;
 KM diagnosis; herpes B virus; ss.
 XX
 OS Herpes simian monkey SAb virus.
 XX
 FH Key Location/Qualifiers
 FT CDS 87..2744
 FT /*tag= a
 FT /product= gB_glycoprotein
 XX
 PN US5487969-A.
 XX
 PD 30-JAN-1996.
 XX
 PF 01-APR-1993; 93US-0042747.
 XX
 PR 01-APR-1993; 93US-0042747.
 XX
 PA (SWBI-) SOUTHWEST FOUND BIOMEDICAL RES.
 XX
 PI Black D, Eberle R, Hilliard J, Scintariello F;
 XX
 DR WPI; 1996-105220/11.
 DR P-PSDB; AAR92747.
 XX
 PT Detection of herpes B virus by PCR amplification of sample DNA - to
 PT detect a specific herpes simian monkey B virus DNA segment.
 XX
 PS Disclosure; Column 23-30; 22pp; English.
 XX
 CC This sequence encodes the herpes simian monkey SAb virus gB
 CC glycoprotein. Such sequences have been used to develop synthetic
 CC DNA primers which have homologous sequences of conserved regions
 CC which flank a divergent region of the gB glycoprotein gene. In
 CC addition, proteins such as the gB glycoprotein have potential use in
 CC the development of serological immunoassays. One approach is to
 CC synthesize peptides which, based on the properties of the predicted
 CC protein sequence, are likely to be immunologically active. Such
 CC peptides can be used as substrate antigens in immunoassays to detect
 CC serum antibodies which recognize this specific peptide sequence.
 CC Synthetic peptides may also be used to produce antibodies against
 CC specific regions of the gB glycoprotein which are unique to one
 CC virus. These can then be used to develop virus-specific
 CC immunoassays for differentiation of SAb virus from other primate

CC alpha-herpes viruses and for identification of antibodies directed
 CC against SAb virus in primate serum samples.
 XX
 SQ Sequence 2943 BP; 462 A; 1103 C; 983 G; 395 T; 0 other;

Query Match 100.0%; Score 18; DB 17; Length 2943;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCAGGCCGCTGTCCTTG 18
 |||||
 Db 2140 AGCAGGCCGCTGTCCTTG 2123

RESULT 9
 AAN60195/c
 ID AAN60195 standard; DNA; 3098 BP.

AC AAN60195;
 XX
 XX 25-MAR-2003 (updated)
 DT 31-JUL-1991 (first entry)

DE Herpes simplex virus glycoprotein gB gene.

XX HSV; gB glycoprotein; vaccine; ss.

XX Herpes simplex virus.

XX Key Location/Qualifiers
 FT CDS 257..296
 /*tag= a

PN EPI170169-A.

PD 05-FEB-1986.

PF 19-JUL-1985; 85EP-0109042.

PR 20-JUL-1984; 84JP-0151766.
 11-DEC-1984; 84JP-0262465.

XX (KAGA) CHERO SERO THERAPEUTIC RES INST.

PI Nozaki C, Makizumi K, Kino Y, Eto, T, Ohtomo N;

DR WFI; 1986-036935/06.
 P-PSDB; AAP60244.

XX Recombinant DNA containing herpes simplex virus gene or fragment -
 PT useful in transformant yeast for prodn. of high purity herpes
 PT simplex virus glycoproteins for use in protective vaccines.

XX Disclosure; Fig. 6; 53pp; English.

XX The gene encodes a gB glycoprotein which is useful for the
 CC production of vaccines conferring protection against herpes simplex
 CC virus infections. The vaccine is safe in use and does not contain a
 CC carcinogenic gene, as with prior vaccines.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 3098 BP; 586 A; 1070 C; 959 G; 483 T; 0 other;

Query Match 100.0%; Score 18; DB 7; Length 3098;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCAGGCCGCTGTCCTTG 18
 |||||
 Db 2352 AGCAGGCCGCTGTCCTTG 2335

RESULT 10
 AAN70681/c
 ID AAN70681 standard; DNA; 3461 BP.

AC AAN70681;

XX 25-MAR-2003 (updated)

DT 10-MAR-2003 (updated)

DT 30-APR-1991 (first entry)

DE Recombinant herpes simplex virus gene.

XX Herpes simplex virus; E.coli; plasmid; vector; promoter; antigen;

XX vaccine; glycoprotein; HSV gB; ss.

XX Herpes simplex virus.

XX Key Location/Qualifiers
 FT CDS 254..2962
 /*tag= a
 /*product= HSV_gB

PN JPE2115288-A.

PD 26-MAY-1987.

PF 14-NOV-1985; 85JP-0255971.

PR 14-NOV-1985; 85JP-0255971.

XX (KAGA) KAGAKU OYOBI KESSEI RYOHU.

DR WFI; 1987-183377/26.
 P-PSDB; AAP70426.

XX Recombinant DNA contg. herpes simplex virus gene - is integrated in
 FT E.coli plasmid vector, contg. incorporated promoter region for
 FT animal cells, giving antigen.

XX Disclosure; Fig 3(1-3); 13pp; Japanese.

XX The gene encoding HSV gB is cloned into E.coli. To increase the
 CC efficiency of expression, the gene encoding the C-terminal region
 CC of gB, which is very hydrophobic, is removed and the remaining gene
 CC is cloned into an expression vector. The vector is composed of pBR322.

CC SV40 early promoter and the HSV gB gene.

CC The construct is useful for the production of a good antigen for
 CC a HSV vaccine.

CC (Updated on 10-MAR-2003 to add missing OS field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 3461 BP; 666 A; 1188 C; 1041 G; 566 T; 0 other;

Query Match 100.0%; Score 18; DB 8; Length 3461;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCAGGCCGCTGTCCTTG 18
 |||||
 Db 2349 AGCAGGCCGCTGTCCTTG 2332

RESULT 11

AAQ14478/c
 ID AAQ14478 standard; DNA; 3465 BP.

AC AAQ14478;

XX 25-MAR-2003 (updated)

DT 23-JAN-1992 (first entry)

DE HSVgB gene.

```

KW Vaccine; antigen; ss.
XX
XX Herpes simplex virus.
OS
XX
XX Key Location/Qualifiers
XX CDS 251..295
XX /*tag= a
XX
XX JP03218397-A.
XX
XX 25-SEP-1991.
XX
XX 21-JUN-1990; 90JP-0161448.
XX
XX 30-NOV-1989; 89JP-0308941.
XX 22-JUN-1989; 89JP-0158238.
XX 21-JUN-1990; 90JP-0161448.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX WPI, 1991-328397/45.
XX P-PDB; AAR14665.
XX
XX HSVB polypeptide(s) obt'd. by recombinant DNA techniques -
XX useful as vaccines against HSV and in diagnosis, can be produced
XX cheaply and safely.
XX
XX Claim 3, Fig 1, 24pp; Japanese.
XX
XX The sequence encodes the HSVB polypeptide. The recombinant protein
XX or esp. truncated forms of it, can be used to prepare vaccines for
XX prophylaxis of HSV infection and for use in diagnostic kits.
XX See also AAQ14479.
XX
XX (updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 3465 BP, 666 A; 1194 C; 1044 G; 561 T; 0 other;
SQ
Query Match 100.0%; Score 18; DB 12; Length 3465;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGCAGGCCGCTGTCTCTTG 18
DB 2349 AGCAGGCCGCTGTCTCTTG 2332
RESULT 12
AAT93650/C
ID AAT93650 standard; DNA; 3471 BP.
XX
XX AAT93650;
XX
XX 26-FEB-1998 (first entry)
XX
XX Herpes simplex virus type 1 glycoprotein B encoding DNA.
XX
XX Herpes simplex virus; HSV; type I; type II; glycoprotein B; vaccine;
XX membrane protein; virus-specific glycoprotein;
XX transmembrane anchor region; ss.
XX
XX Herpes simplex virus type 1.
XX
XX Key Location/Qualifiers
XX PH CMT_1signal 55..62
XX FT /*tag= a
XX FT /*note= "putative"
XX FT TATA_signal 125..131
XX FT /*tag= b
XX FT /*note= "putative"
XX CDS 309..3023
XX FT /*tag= c
XX FT 3166..3173
XX FT /*tag= d
XX

```

```

FT /*note= "putative"
FT polyA_signal 3409..3416
FT /*tag= e
FT /*note= "putative"
XX
XX US5648079-A.
XX
XX 15-JUL-1997.
XX
XX 08-DEC-1994; 94US-0351875.
XX
XX 20-OCT-1986; 86US-0921730.
XX 06-APR-1984; 84US-0597784.
XX 20-SEP-1990; 90US-0587179.
XX 21-DEC-1992; 92US-0993415.
XX 18-OCT-1993; 93US-0138717.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Burke RL, Pachl C, Valenzuela PDT;
XX
XX WPI, 1997-372022/34.
XX P-PDB; AAW34552.
XX
XX Vaccines against herpes simplex virus infection - containing
XX recombinant HSV glycoprotein B
XX
XX Disclosure: Fig 4A-F; 33pp; English.
XX
XX The present sequence is from the Herpes simplex virus (HSV) type I
XX strain Patton and encodes a glycoprotein B. HSV is a double stranded
XX virus packaged within an icosahedral nucleocapsid enveloped within a
XX membrane. The membrane includes a number of virus-specific
XX glycoproteins, with glycoprotein B being one of the most abundant.
XX Glycoprotein B from both HSV type I and type II are cross reactive. New
XX vaccines comprising recombinantly produced glycosylated glycoprotein B
XX that has a deletion of at least part of the transmembrane anchor region,
XX in combination with a carrier and an adjuvant have been produced. The
XX vaccines are used for immunising humans against HSV (HSV type 1 or
XX HSV type 2) by vaccination before or after primary infection with HSV.
XX
XX Sequence 3471 BP, 639 A; 1201 C; 1069 G; 562 T; 0 other;
SQ
Query Match 100.0%; Score 18; DB 18; Length 3471;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGCAGGCCGCTGTCTCTTG 18
DB 2398 AGCAGGCCGCTGTCTCTTG 2381
RESULT 13
AAN80907/C
ID AAN80907 standard; DNA; 3472 BP.
XX
XX AAN80907;
XX
XX 25-MAR-2003 (updated)
XX 09-MAR-1992 (first entry)
XX
XX Sequence of Herpes simplex virus (HSV) glycoprotein B (gB) 1 DNA.
XX
XX Vaccine; herpes simplex virus; therapy; ss.
XX
XX Herpes simplex virus.
XX
XX Key Location/Qualifiers
XX PH CDS 309..3023
XX FT /*tag= a
XX
XX MO8802634-A.
XX

```

PD 21-APR-1988.
 XX 20-OCT-1987; 87WO-US02709.
 XX 20-JUN-1987; 87US-0079605.
 PR 20-OCT-1986; 86US-0921213.
 XX (CHIR) CHIRON CORP.
 XX Burke LR, Pachl C, Valenzuela P;
 DR WPI, 1986-119368/17.
 DR P-PSDB; AAP80914.
 XX
 PT Vaccine for treatment of herpes simplex virus - contains
 PT recombinant HSV glyco:proteins B and D
 XX
 PS Example; Fig 4; 71pp; English.
 CC Prepn. of recombinant gB and gD is described in WO88504587. The
 CC amino acid sequence and DNA sequence for gB 1 presented in AAP80914
 CC and AAN80907 differ from that originally presented in Table 1 of
 CC International Publication No. WO 85/04587. The DNA sequence in Table
 CC 1 contains an error in that an additional nucleotide (G) is listed
 CC at position 607 which resulted in a shift in reading frame relative
 CC to AAN80907 from which this nucleotide has been deleted.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 SQ Sequence 3472 BP; 639 A; 1202 C; 1068 G; 563 T; 0 other;
 Query Match 100.0%; Score 18; DB 9; Length 3472;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGCAGCGCGCTGCTCTTG 18
 DB 2398 AGCAGCGCGCTGCTCTTG 2381

RESULT 14
 AAQ48496/C
 ID AAQ48496 standard; DNA; 3472 BP.
 XX
 AC AAQ48496;
 XX
 DT 25-MAR-2003 (updated)
 DT 29-MAR-1994 (first entry)
 XX
 DE Glycoprotein B (gB1).
 XX
 KW Glycoprotein; gB1; gB2; herpes simplex virus; HSV-1; HSV-2;
 KW flanking; initiation; termination; transcription; translation;
 KW regulatory sequence; ss.
 XX
 OS Herpes simplex virus 1 strain Patton.
 XX
 FH Key Location/Qualifiers
 FT CDS 309..3023
 FT /*tag= a
 FT
 XX
 PN US5244792-A.
 XX
 PD 14-SEP-1993.
 XX
 PF 20-SEP-1990; 90US-0587179.
 XX
 PR 06-APR-1984; 84US-0597784.
 PR 20-OCT-1986; 86US-0921730.
 PR 20-SEP-1990; 90US-0587179.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Burke RL, Pachl C, Valenzuela PDT;

XX WPI; 1993-302641/38.
 DR P-PSDB; AAR41778.
 XX
 PT DNA construct for expressing HSV glycosylated polypeptide -
 PT useful for vaccinating against HSV-1 and -2 infections
 XX
 PS Disclosure; Fig 4; 33pp; English.
 CC
 CC DNA constructs for expressing a glycosylated polypeptide in a
 CC eukaryotic host cell comprises (1) an oligonucleotide sequence (OS)
 CC free of natural flanking sequences, encoding glycoprotein B (gB)
 CC of HSV, or its fragments, and (2) 5' initiating and 3' terminating
 CC transcriptional and translational regulatory sequences flanking OS,
 CC at least one of these sequences not being from HSV. The HSV-1 and
 CC HSV-2 sequences are given in AAQ48496 and AAQ48497 respectively.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC
 SQ Sequence 3472 BP; 639 A; 1202 C; 1069 G; 562 T; 0 other;
 Query Match 100.0%; Score 18; DB 14; Length 3472;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGCAGCGCGCTGCTCTTG 18
 DB 2398 AGCAGCGCGCTGCTCTTG 2381

RESULT 15
 AAQ48497/C
 ID AAQ48497 standard; DNA; 3472 BP.
 XX
 AC AAQ48497;
 XX
 DT 25-MAR-2003 (updated)
 DT 29-MAR-1994 (first entry)
 XX
 DE Glycoprotein B (gB2).
 XX
 KW Glycoprotein; gB1; gB2; herpes simplex virus; HSV-1; HSV-2;
 KW flanking; initiation; termination; transcription; translation;
 KW regulatory sequence; ss.
 XX
 OS Herpes simplex virus 2 strain 333.
 XX
 FH Key Location/Qualifiers
 FT CDS 309..3023
 FT /*tag= a
 FT
 XX
 PN US5244792-A.
 XX
 PD 14-SEP-1993.
 XX
 PF 20-SEP-1990; 90US-0587179.
 XX
 PR 06-APR-1984; 84US-0597784.
 PR 20-OCT-1986; 86US-0921730.
 PR 20-SEP-1990; 90US-0587179.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Burke RL, Pachl C, Valenzuela PDT;
 XX
 DR WPI; 1993-302641/38.
 DR P-PSDB; AAR41779.
 XX
 PT DNA construct for expressing HSV glycosylated polypeptide -
 PT useful for vaccinating against HSV-1 and -2 infections
 XX
 PS Disclosure; Fig 4; 33pp; English.
 XX
 CC DNA constructs for expressing a glycosylated polypeptide in a

CC eukaryotic host cell sequences (1) an oligonucleotide sequence (OS)
CC fire of natural flanking sequences, encoding glycoprotein B (GB)
CC of HSV, or its fragments and (2) 5' initiating and 3' terminating
CC transcriptional and translational regulatory sequences flanking OS
CC at least one of these sequences not being from HSV. The HSV-1 and
CC HSV-2 sequences are given in MA04896 and MA04897 respectively.
CC (Updated on 25-MAR-2003 to correct Pf field.)

50 Sequence 3472 BP; 661 A; 1193 C; 1053 G; 565 T; 0 other;

Query Match	Score	DB	Length
100.0%	18	14	3472

Matches 18; Conservative 0; Mismatches

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 AGCAGGCCGCTGTCCTG 18
          |||||
Db      . 2407 AGCAGGCCGCTGTCCTG 2390

```

Db , 2407 AGCAGGCCGCTGTCCTTG 2390

Search completed: September 16, 2003, 18:36:58
Job time : 91.1782 secs

Job time : 91.1782 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 18:24:35 ; Search time 22.9901 Seconds
(without alignments)
345.579 Million cell updates/sec

Title: US-09-594-065-2

Perfect score: 18

Sequence: 1 agcagccgcgtccttg 18

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues 1139956

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	2713	US-08-804-439A-13	Sequence 13, Appl
2	18	100.0	2713	US-08-720-229-13	Sequence 13, Appl
3	18	100.0	2943	US-08-042-747A-7	Sequence 7, Appl
4	18	100.0	3472	5244792-1	Patent No. 5244792
5	18	100.0	5244792-2	5244792-2	Patent No. 5244792
6	17	94.4	855	US-09-252-991A-1201	Sequence 1201, Ap
7	17	94.4	1491	US-09-252-991A-1243	Sequence 1243, Ap
8	17	94.4	1551	US-09-252-991A-1362	Sequence 1362, Ap
9	17	94.4	6244	US-08-076-726-15	Sequence 15, Appl
10	17	94.4	6244	US-08-260-452-8	Sequence 8, Appl
11	17	94.4	6244	US-08-481-970-8	Sequence 8, Appl
12	17	94.4	6244	US-08-897-718-8	Sequence 8, Appl
13	17	94.4	6244	US-09-163-269-8	Sequence 8, Appl
14	17	94.4	711	US-09-252-991A-12195	Sequence 12195, A
15	16.4	91.1	1110	US-08-851-088-9	Sequence 9, Appl
16	16.4	91.1	1149	US-09-252-991A-12362	Sequence 12362, A
17	16.4	91.1	1251	US-09-252-991A-12303	Sequence 12303, A
18	16.4	91.1	1278	US-09-345-468-24	Sequence 24, Appl
19	16.4	91.1	1896	US-09-414-453A-24	Sequence 24, Appl
20	16.4	91.1	1896	US-09-832-312-24	Sequence 24, Appl
21	16.4	91.1	2170	US-09-345-468-11	Sequence 11, Appl
22	16.4	91.1	2170	US-09-414-453A-11	Sequence 11, Appl
23	16.4	91.1	2170	US-09-832-312-11	Sequence 11, Appl
24	16.4	91.1	2170	US-09-832-312-11	Sequence 11, Appl
25	16.4	91.1	2194	US-09-310-463-19	Sequence 19, Appl
26	16.4	91.1	2194	US-08-842-248A-19	Sequence 19, Appl
27	16.4	91.1	2194	US-08-842-248A-19	Sequence 19, Appl

C 28	16.4	91.1	3177	1	US-08-042-747A-4	Sequence 4, Appl
C 29	16.4	91.1	4144	3	US-08-851-089-12	Sequence 12, Appl
C 30	16.4	91.1	4144	3	US-08-851-089-13	Sequence 13, Appl
C 31	16	88.9	2244	1	US-08-476-519-10	Sequence 10, Appl
C 32	16	88.9	2244	5	PCT-US95-09323-10	Sequence 10, Appl
C 33	16	88.9	2334	1	US-08-476-519-1	Sequence 1, Appl
C 34	16	88.9	2334	5	PCT-US95-09323-1	Sequence 1, Appl
C 35	15.4	85.6	3014	1	US-08-629-939-1	Sequence 1, Appl
C 36	15.4	85.6	3014	4	US-08-759-873-1	Sequence 1, Appl
C 37	15.4	85.6	3014	4	US-09-016-434-1244	Sequence 1244, Ap
C 38	15.4	85.6	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C 39	15.4	85.6	4411529	3	US-09-103-840A-1	Sequence 1, Appl
C 40	14.8	82.2	290	4	US-09-313-294A-6122	Sequence 6122, Ap
C 41	14.8	82.2	471	4	US-09-252-991A-263	Sequence 263, App
C 42	14.8	82.2	516	4	US-09-252-991A-15658	Sequence 15658, A
C 43	14.8	82.2	684	4	US-09-252-991A-11110	Sequence 11110, A
C 44	14.8	82.2	747	3	US-08-998-416-246	Sequence 246, App
C 45	14.8	82.2	765	4	US-09-199-637A-324	Sequence 324, App

ALIGNMENTS

RESULT 1
US-08-804-439A-13/c
Sequence 13, Application US/08804439A
Patent No. 6015565
GENERAL INFORMATION:
APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Martin L.
TITLE OF INVENTION: GLYCOPROTEIN B OF THE REVU/KSHV
TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESSES:
ADDRESS: Fish & Richardson P.C.
STREET: 4225 Executive Square, Ste 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,439A
FILING DATE: February 21, 1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09176/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
TELEX:
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2713 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-804-439A-13
Query Match 100.0%; Score 18; DB 3; Length 2713;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGCAGCGCGTCTCTTG 18
DB 2096 AGCAGCGCGTCTCTTG 2079

RESULT 2
US-08-720-229-13/C
Sequence 13, Application US/08720229
Patent No. 6022542
GENERAL INFORMATION:
APPLICANT: Rose, Timothy M.
APPLICANT: Boesch, Marnix L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RRV/KSHV
TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,229
FILING DATE: 26-SEP-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schiffe, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2713 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-720-229-13

Query Match 100.0%; Score 18; DB 3; Length 2713;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCAGGCCGCTGCTCTTG 18
|||||
Db 2096 AGCAGGCCGCTGCTCTTG 2079

RESULT 3
US-08-042-747A-7/C
Sequence 7, Application US/08042747A
Patent No. 5487969
GENERAL INFORMATION:
APPLICANT: Eberle, Richard
APPLICANT: Black, Darla
APPLICANT: Scimicciello, Franco
APPLICANT: Hilliard, Julia K.
TITLE OF INVENTION: Cloning and Amplification of Monkey B
TITLE OF INVENTION: Virus Genes
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cox & Smith Incorporated
STREET: 112 East Pecan Street, Suite 2000
CITY: San Antonio
STATE: Texas
COUNTRY: USA
ZIP: 78205

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/042,747A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haymond, M. Bradley
REGISTRATION NUMBER: 35186
REFERENCE/DOCKET NUMBER: S-0072.179
TELECOMMUNICATION INFORMATION:
TELEPHONE: 210-554-5500
TELEFAX: 210-226-8395
TELEX: 767609
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2943 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 87..2744
US-08-042-747A-7

Query Match 100.0%; Score 18; DB 1; Length 2943;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCAGGCCGCTGCTCTTG 18
|||||
Db 2140 AGCAGGCCGCTGCTCTTG 2123

RESULT 4
5244792-1/C
Patent No. 5244792
APPLICANT: BURKE, RAE L.; PACHL, CAROL; VALENZUELA, PABLO D.T.
TITLE OF INVENTION: EXPRESSION OF RECOMBINANT GLYOPROTEIN
B FROM HERPES SIMPLEX VIRUS
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587,179
FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 921,730
FILING DATE: 20-OCT-1986
APPLICATION NUMBER: 597,784
FILING DATE: 06-APR-1984
SEQ ID NO: 1:
LENGTH: 3472
5244792-1

Query Match 100.0%; Score 18; DB 6; Length 3472;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCAGGCCGCTGCTCTTG 18
|||||
Db 2398 AGCAGGCCGCTGCTCTTG 2381

RESULT 5
5244792-2/C
Patent No. 5244792
APPLICANT: BURKE, RAE L.; PACHL, CAROL; VALENZUELA, PABLO D.T.
TITLE OF INVENTION: EXPRESSION OF RECOMBINANT GLYOPROTEIN
B FROM HERPES SIMPLEX VIRUS
NUMBER OF SEQUENCES: 19


```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,179
; FILING DATE: 20-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 921,730
; FILING DATE: 20-OCT-1986
; APPLICATION NUMBER: 597,784
; FILING DATE: 06-APR-1984
; SEQ ID NO:2:
; LENGTH: 3472
5244792-2

Query Match          100.0%; Score 18; DB 6; Length 3472;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGCAGCCGCTGTCTTG 18
Db 2407 AGCAGCCGCTGTCTTG 2390

RESULT 6
US-09-252-991A-1201/c
; Sequence 1201, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1201
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1201

Query Match          94.4%; Score 17; DB 4; Length 855;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GCAGCCGCTGTCTTG 18
Db 539 GCAGCCGCTGTCTTG 523

RESULT 7
US-09-252-991A-1243/c
; Sequence 1243, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1243
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1243

Query Match          94.4%; Score 17; DB 4; Length 1491;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GCAGCCGCTGTCTTG 18
Db 424 GCAGCCGCTGTCTTG 408

RESULT 8
US-09-252-991A-1362
; Sequence 1362, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1362
; LENGTH: 1551
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1362

Query Match          94.4%; Score 17; DB 4; Length 1551;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GCAGCCGCTGTCTTG 18
Db 1116 GCAGCCGCTGTCTTG 1132

RESULT 9
US-08-076-726-15/c
; Sequence 15, Application US/08076726
; Patent No. 5464758
; GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; TITLE OF INVENTION: Tight Control of Gene Expression in
; TITLE OF INVENTION: Eucaryotic Cells by Tetracycline-responsive Promoters
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: United States of America
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/076,726
; FILING DATE: 14-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Edmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942, 2490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
```

TELEFAX: (202) 371-2540
TELEX: 248636 SSK
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 6244 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-076-726-15

Query Match 94.4%; Score 17; DB 1; Length 6244;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGTCTT 17
|||||
DB 991 AGCAGCGCGCTGTCTT 975

RESULT 10

US-08-260-452-8/c
Sequence 8, Application US/08260452
Patent No. 5650298
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
APPLICANT: Bujard, Hermann
APPLICANT: Salfeld, Jochen
APPLICANT: Voss, Jeffrey
TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic
TITLE OF INVENTION: Cells by Tetracycline-responsive Promoters
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/260,452
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076,327
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6244 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomagalovirus
STRAIN: Towne (hCMV)
IMMEDIATE SOURCE:
CLONE: pUHD BGR3
US-08-260-452-8

Query Match 94.4%; Score 17; DB 1; Length 6244;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGTCTT 17
|||||
DB 991 AGCAGCGCGCTGTCTT 975

RESULT 11

US-08-481-970-8/c
Sequence 8, Application US/08481970
Patent No. 5859310
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
APPLICANT: Bujard, Hermann
APPLICANT: Salfeld, Jochen
APPLICANT: Voss, Jeffrey
TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,970
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/260,452
FILING DATE: 14-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076,327
FILING DATE: 14-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6244 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomagalovirus
STRAIN: Towne (hCMV)
IMMEDIATE SOURCE:
CLONE: pUHD BGR3
US-08-481-970-8

Query Match 94.4%; Score 17; DB 2; Length 6244;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGTCTT 17
|||||
DB 991 AGCAGCGCGCTGTCTT 975

RESULT 12

US-08-897-719-8/c
Sequence 8, Application US/08897719
Patent No. 5922927

```

GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
APPLICANT: Bujard, Hermann
APPLICANT: Salfeld, Jochen
APPLICANT: Voss, Jeffrey
TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic
CELLS BY TETRACYCLINE-RESPONSIVE PROMOTERS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Lahnive & Cockfield
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,719
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/260,452
FILING DATE:
APPLICATION NUMBER: 08/076,327
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6244 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
STRAIN: Towne (hcmv)
IMMEDIATE SOURCE:
CLONE: PUHD BGR3
US-08-897-719-8

Query Match 94.4% Score 17; DB 2; Length 6244;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGCAGCGCGCTGCTCTT 17
DB 991 AGCAGCGCGCTGCTCTT 975

RESULT 13
US-09-163-269-8/c
Sequence 8, Application US/09163269
Patent No. 6252136
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
APPLICANT: Bujard, Hermann
APPLICANT: Salfeld, Jochen
APPLICANT: Voss, Jeffrey
TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled
Transcript
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Lahnive & Cockfield
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/163,269
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/481,970
FILING DATE:
APPLICATION NUMBER: 08/076,327
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6244 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
STRAIN: Towne (hcmv)
IMMEDIATE SOURCE:
CLONE: PUHD BGR3
US-09-163-269-8

Query Match 94.4% Score 17; DB 3; Length 6244;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGCAGCGCGCTGCTCTT 17
DB 991 AGCAGCGCGCTGCTCTT 975

RESULT 14
US-09-252-991A-12195/c
Sequence 12195, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 12195
LENGTH: 711
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12195

Query Match 91.1% Score 16.4; DB 4; Length 711;
Best Local Similarity 94.4%; Pred. No. 36;
```

```

CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/163,269
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/481,970
FILING DATE:
APPLICATION NUMBER: 08/076,327
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6244 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
STRAIN: Towne (hcmv)
IMMEDIATE SOURCE:
CLONE: PUHD BGR3
US-09-163-269-8

Query Match 94.4% Score 17; DB 3; Length 6244;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGCAGCGCGCTGCTCTT 17
DB 991 AGCAGCGCGCTGCTCTT 975

RESULT 14
US-09-252-991A-12195/c
Sequence 12195, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 12195
LENGTH: 711
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12195

Query Match 91.1% Score 16.4; DB 4; Length 711;
Best Local Similarity 94.4%; Pred. No. 36;
```

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AGCAGGCCGCTGCTTG 18
 |||||
 DB 514 AGCAGGCCGCTGCTTG 497

RESULT 15
 US-08-851-088-9
 / Sequence 9, Application US/08851088
 / Patent No. 5952208
 / GENERAL INFORMATION:
 / APPLICANT: Darzine, Aldis
 / APPLICANT: X1, Lei
 / APPLICANT: Childs, John D.
 / APPLICANT: Monticello, Daniel J.
 / APPLICANT: Squires, Charles H.
 / TITLE OF INVENTION: DSZ Gene Expression In Pseudomonas Hosts
 / NUMBER OF SEQUENCES: 23
 / CORRESPONDENCE ADDRESS:
 / ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.
 / STREET: Two Militia Drive
 / CITY: Lexington
 / STATE: Massachusetts
 / COUNTRY: USA
 / ZIP: 02173
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patent In Release #1.0, Version #1.30
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/851,088
 / FILING DATE: 05-May-1997
 / CLASSIFICATION: 435
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 08/835,185
 / FILING DATE: 07-APR-1997
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Elmore, Carolyn S.
 / REGISTRATION NUMBER: 37,567
 / REFERENCE/DOCKET NUMBER: EBC96-06A
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (781) 861-6240
 / TELEFAX: (781) 861-9540
 / INFORMATION FOR SEQ. ID NO: 9:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 110 base pairs
 / TYPE: nucleic acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / MOLECULE TYPE: DNA (genomic)
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 1..1107
 / US-08-851-088-9

Query Match 91.1%; Score 16.4; DB 2; Length 1110;
 Best Local Similarity 94.4%; Pred. No. 36;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AGCAGGCCGCTGCTTG 18
 |||||
 DB 620 AGCAGGCCGCTGCTTG 637

Search completed: September 16, 2003, 20:42:36
 Job time: 27.9901 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 19:14:25 ; Search time 65.0495 Seconds
(without alignments)
680.690 Million cell updates/sec

Title: US-09-594-065-2

Perfect score: 18
Sequence: 1 agcagggccgctgccttg 18

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1660708 seqs, 122959015 residues

Total number of hits satisfying chosen parameters: 3321416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_NEW_PUB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	154746	12	US-09-827-688-8
2	17	94.4	1431	9	US-09-815-242-7812
3	17	94.4	6244	9	US-09-281-674-8
4	17	94.4	6244	10	US-09-892-227-8
5	16.4	91.1	609	13	US-10-027-632-12963
6	16.4	91.1	609	13	US-10-027-632-12964
7	16.4	91.1	1896	9	US-09-832-312-24
8	16.4	91.1	2061	10	US-09-954-456-960
9	16.4	91.1	2170	9	US-09-832-312-11
10	16.4	91.1	2194	12	US-10-143-618-19
11	16.4	91.1	2194	14	US-10-139-662-19
12	16.4	91.1	2194	14	US-10-139-683-19
13	16.4	91.1	13553	11	US-09-764-891-7432
14	16	88.9	976	14	US-10-198-846-11382
15	15.4	85.6	489	10	US-09-764-868-222
16	15.4	85.6	573	13	US-10-027-632-15146

C 17	15.4	85.6	622	13	US-10-027-632-267576	Sequence 267576,
C 18	15.4	85.6	2190	14	US-10-156-761-2735	Sequence 2735, Ap
C 19	15.4	85.6	2453	10	US-09-974-298-111	Sequence 111, Ap
C 20	15.4	85.6	2470	9	US-09-870-937-9	Sequence 9, Appli
C 21	15.4	85.6	2470	10	US-09-954-456-1121	Sequence 1121, Ap
C 22	15.4	85.6	2470	10	US-09-880-107-1725	Sequence 1725, Ap
C 23	15.4	85.6	2470	12	US-10-354-358-7	Sequence 7, Appli
C 24	15.4	85.6	2519	13	US-10-071-766-82	Sequence 82, Appli
C 25	15.4	85.6	2799	9	US-09-935-368-2	Sequence 2, Appli
C 26	15.4	85.6	5003	14	US-10-207-655-62	Sequence 62, Appli
C 27	15.4	85.6	23098	12	US-10-017-161-2005	Sequence 2005, Ap
C 28	15.4	85.6	902568	14	US-10-156-761-1	Sequence 1, Appli
C 29	15	83.3	262	9	US-09-250-883-1	Sequence 1, Appli
C 30	15	83.3	263	9	US-09-250-883-2	Sequence 2, Appli
C 31	15	83.3	1332	9	US-09-250-883-14	Sequence 14, Appli
C 32	15	83.3	1929	14	US-10-106-698-805	Sequence 805, App
C 33	15	83.3	3446	12	US-09-769-734-4	Sequence 4, Appli
C 34	14.8	82.2	60	12	US-09-908-975-13289	Sequence 13289, A
C 35	14.8	82.2	262	9	US-09-923-876-2088	Sequence 2088, Ap
C 36	14.8	82.2	371	10	US-09-796-692-6249	Sequence 6249, Ap
C 37	14.8	82.2	371	14	US-10-040-862-6249	Sequence 6249, Ap
C 38	14.8	82.2	424	13	US-10-001-857-58	Sequence 58, Appli
C 39	14.8	82.2	439	10	US-09-796-692-2556	Sequence 2556, Ap
C 40	14.8	82.2	439	14	US-10-040-862-2556	Sequence 2556, Ap
C 41	14.8	82.2	471	11	US-09-918-995-32269	Sequence 32269, A
C 42	14.8	82.2	478	13	US-10-027-632-287388	Sequence 287388,
C 43	14.8	82.2	618	12	US-10-308-779-67	Sequence 67, Appli
C 44	14.8	82.2	765	11	US-09-975-719-324	Sequence 324, App
C 45	14.8	82.2	810	14	US-10-156-761-2123	Sequence 2123, Ap

ALIGNMENTS

RESULT 1
US-09-827-688-8
; Sequence 8, Application US/09827688
; Publication No. US20030165476A1
; GENERAL INFORMATION:
; APPLICANT: ORSON, FRANK
; APPLICANT: KINSEY, BERMA
; APPLICANT: BHOAL, BALBIR
; TITLE OF INVENTION: MAGROGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION D
; FILE REFERENCE: P01949051/10004014
; CURRENT APPLICATION NUMBER: US/09/827,688
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/195,680
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 154746
; TYPE: DNA
; ORGANISM: HERPESVIRUS 2
US-09-827-688-8

Query Match 100.0%; Score 18; DB 12; Length 154746;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCAGGCGCGTGCCTTG 18
DB 54028 AGCAGGCGCGTGCCTTG 54045

RESULT 2
US-09-815-242-7812/c
; Sequence 7812, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.

APPLICANT: Zykand, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21 078
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7812
LENGTH: 1431
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1431)
US-09-815-242-7812

Query Match 94.4%; Score 17; DB 9; Length 1431;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GCAGGCCGCTGCTCTTG 18
DB 364 GCAGGCCGCTGCTCTTG 348

RESULT 3
US-09-281-674-8/c
Sequence 8, Application US/09281674
Patent No. US2002007307A1
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
Bujard, Hermann
Salfeld, Jochen
Voess, Jeffrey
TITLE OF INVENTION: Methods for Regulating Gene Expression
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/281,674
FILING DATE: 30-Mar-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/479,306
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/260,452
FILING DATE: 14-JUN-1994
APPLICATION NUMBER: 08/076,327
FILING DATE: 14-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6244 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
STRAIN: Towne (hCMV)
IMMEDIATE SOURCE:
CLONE: DHD BGR3
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-281-674-8

Query Match 94.4%; Score 17; DB 9; Length 6244;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCAGGCCGCTGCTCTT 17
DB 991 AGCAGGCCGCTGCTCTT 975

RESULT 4
US-09-892-227-8/c
Sequence 8, Application US/09892227
Patent No. US20020152487A1
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
Bujard, Hermann
Salfeld, Jochen
Voess, Jeffrey
TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled
Transcription
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/892,227
FILING DATE: 25-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/163,269
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/076,327
FILING DATE: 14-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503

REFERENCE/DOCKET NUMBER: BBI-013CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6244 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
STRAIN: Towne (hcmv)
IMMEDIATE SOURCE:
CLONE: PUHD BGR3
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-892-227-8

Query Match 94.4%; Score 17; DB 10; Length 6244;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGTCTT 17
|||||
Db 991 AGCAGCGCGCTGTCTT 975

RESULT 5
US-10-027-632-12963
Sequence 12963, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 12963
LENGTH: 609
TYPE: DNA
ORGANISM: Human
US-10-027-632-12963

Query Match 91.1%; Score 16.4; DB 13; Length 609;
Best Local Similarity 94.4%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGTCTTG 18
|||||
Db 97 AGCAGCGCGCTGTCTTG 114

RESULT 6
US-10-027-632-12964
Sequence 12964, Application US/10027632
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 12964
LENGTH: 609
TYPE: DNA
ORGANISM: Human
US-10-027-632-12964

Query Match 91.1%; Score 16.4; DB 13; Length 609;
Best Local Similarity 94.4%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGTCTTG 18
|||||
Db 97 AGCAGCGCGCTGTCTTG 114

RESULT 7
US-09-832-312-24/c
Sequence 24, Application US/09832312
Patent No. US20010049829A1
GENERAL INFORMATION:
APPLICANT: Busfield et al.
TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
FILE REFERENCE: 7853-234
CURRENT APPLICATION NUMBER: US/09/832,312
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: 09/610,118
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/503,387
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 09/454,824
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 09/345,468
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 78
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 24
LENGTH: 1896
TYPE: DNA
ORGANISM: Homo sapiens
US-09-832-312-24

Query Match 91.1%; Score 16.4; DB 9; Length 1896;
Best Local Similarity 94.4%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGTCTTG 18
|||||
Db 1496 AGCAGCGCGCTGTCTTG 1479

RESULT 8

US-09-954-456-960/C
; Sequence 960, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIORITY FILING DATE: 2001-09-18
; PRIORITY APPLICATION NUMBER: US/60/233,617
; PRIORITY FILING DATE: 2000-09-18
; PRIORITY APPLICATION NUMBER: US/60/234,052
; PRIORITY FILING DATE: 2000-09-20
; PRIORITY APPLICATION NUMBER: US/60/234,923
; PRIORITY FILING DATE: 2000-09-25
; PRIORITY APPLICATION NUMBER: US/60/235,134
; PRIORITY FILING DATE: 2000-09-25
; PRIORITY APPLICATION NUMBER: US/60/235,637
; PRIORITY FILING DATE: 2000-09-26
; PRIORITY APPLICATION NUMBER: US/60/235,638
; PRIORITY FILING DATE: 2000-09-26
; PRIORITY APPLICATION NUMBER: US/60/235,711
; PRIORITY FILING DATE: 2000-09-27
; PRIORITY APPLICATION NUMBER: US/60/235,720
; PRIORITY FILING DATE: 2000-09-27
; PRIORITY APPLICATION NUMBER: US/60/235,840
; PRIORITY FILING DATE: 2000-09-27
; PRIORITY APPLICATION NUMBER: US/60/235,863
; PRIORITY FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 960
; LENGTH: 2061
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-960

Query Match 91.1%; Score 16.4; DB 10; Length 2061;
Best Local Similarity 94.4%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGGCCGCTGTCTTGG 18
|||||
Db 1496 AGCAGGCCGCTGTCTTGG 1479

RESULT 9
US-09-832-312-11/C
; Sequence 11, Application US/09832312
; Patent No. US20010049829A1
; GENERAL INFORMATION:
; APPLICANT: Buefield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; CURRENT FILING DATE: 2001-04-09
; PRIORITY APPLICATION NUMBER: 09/610,118
; PRIORITY FILING DATE: 2000-06-30
; PRIORITY APPLICATION NUMBER: 09/503,387
; PRIORITY FILING DATE: 2000-02-14
; PRIORITY APPLICATION NUMBER: 09/454,824
; PRIORITY FILING DATE: 1999-12-06
; PRIORITY APPLICATION NUMBER: 09/345,468
; PRIORITY FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 2170
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-832-312-11

Query Match 91.1%; Score 16.4; DB 9; Length 2170;
Best Local Similarity 94.4%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGGCCGCTGTCTTGG 18
|||||
Db 1545 AGCAGGCCGCTGTCTTGG 1528

RESULT 10
US-10-143-618-19/C
; Sequence 19, Application US/10143618
; Publication No. US20030166073A1
; GENERAL INFORMATION:
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: Family of Immunoregulators Designated
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Janis C. Henry, Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM/PC Compatible
; OPERATING SYSTEM: Microsoft Word 7.0
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/143,618
; FILING DATE: 08-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/842,248A
; FILING DATE: April 24, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C.
; REGISTRATION NUMBER: 34,147
; REFERENCE/DOCKET NUMBER: 2624
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2194 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: Dpm17
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 67..1959
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-143-618-19

Query Match 91.1%; Score 16.4; DB 12; Length 2194;
Best Local Similarity 94.4%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGGCCGCTGTCTTGG 18
|||||
Db 1562 AGCAGGCCGCTGTCTTGG 1545

RESULT 11
US-10-139-662-19/C
; Sequence 19, Application US/10139662
; Publication No. US20030027358A1


```
; GENERAL INFORMATION:
; APPLICANT: Cosman, David J.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Borges, Luis
; TITLE OF INVENTION: Family of Immunoregulators Designated Leukocyte Immunoglobulin-
; FILE REFERENCE: 2624-A
; CURRENT APPLICATION NUMBER: US/10/139,662
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 09/310,463
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 08/842,248
; PRIOR FILING DATE: 1997-04-24
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 2194
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)..(1962)
US-10-139-662-19

Query Match          91.1%; Score 16.4; DB 14; Length 2194;
Best Local Similarity 94.4%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AGCAGCGCGCTGTCCTTG 18
Db      1562 AGCAGCGCGCTGTCCTTG 1545

RESULT 12
US-10-139-683-19/C
; Sequence 19, Application US/10139683
; Publication No. US20030060614A1
; GENERAL INFORMATION:
; APPLICANT: Cosman, David J.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Borges, Luis
; TITLE OF INVENTION: Family of Immunoregulators Designated Leukocyte Immunoglobulin-
; FILE REFERENCE: 2624-A
; CURRENT APPLICATION NUMBER: US/10/139,683
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 09/310,463
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 08/842,248
; PRIOR FILING DATE: 1997-04-24
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 2194
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)..(1962)
US-10-139-683-19

Query Match          91.1%; Score 16.4; DB 14; Length 2194;
Best Local Similarity 94.4%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
; Publication No. US20030077809A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7432
; LENGTH: 13553
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7432

Query Match          91.1%; Score 16.4; DB 11; Length 13553;
Best Local Similarity 94.4%; Pred. No. 30;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AGCAGCGCGCTGTCCTTG 18
Db      1984 AGCAGCGCTGTCCTTG 1967

RESULT 14
US-10-198-846-11382/C
; Sequence 11382, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11382
; LENGTH: 976
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-11382

Query Match          88.9%; Score 16; DB 14; Length 976;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 CAGCGCGCTGTCCTTG 18
Db      589 CAGCGCGCTGTCCTTG 574

RESULT 15
US-09-764-868-222
; Sequence 222, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 222
```

```

: LENGTH: 489
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (424)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (441)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (449)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-764-866-222

```

```
Query Match      85.6%; Score 15.4; DB 10; Length 489;
Best Local Similarity 94.1%; Pred. No. 1.5e+2;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy	1	AGCAGGCCGCTGTCCTT	17
Db	34	AGCAGGCCCTGTCCTT	50

Search completed: September 16, 2003, 22:54:15
Job time : 71.0495 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 18:22:25 ; Search time 911.049 Seconds
(without alignments)
480.194 Million cell updates/sec

Title: US-09-594-065-2

Perfect score: 1 agcagccgcgcgccttg 18

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estrov:*
6: em_estcp1:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_dln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.4	91.1	195	9	AM663294 bh77a10.y
2	16.4	91.1	287	14	T63918 yc08c12.fl
3	16.4	91.1	308	28	A2081267 SLG2A2 Fe
4	16.4	91.1	356	13	BY310214 BY310214

5	16.4	91.1	364	13	BY664919
6	16.4	91.1 <td>366</td> <td>13 <td>BY064920</td> </td>	366	13 <td>BY064920</td>	BY064920
7	16.4	91.1 <td>373</td> <td>13 <td>BY661558</td> </td>	373	13 <td>BY661558</td>	BY661558
8	16.4	91.1 <td>388</td> <td>13 <td>BY066019</td> </td>	388	13 <td>BY066019</td>	BY066019
9	16.4	91.1 <td>391</td> <td>9</td> <td>AV661443</td>	391	9	AV661443
10	16.4	91.1 <td>424</td> <td>13 <td>BY223280</td> </td>	424	13 <td>BY223280</td>	BY223280
11	16.4	91.1 <td>442</td> <td>13 <td>BQ317016</td> </td>	442	13 <td>BQ317016</td>	BQ317016
12	16.4	91.1 <td>522</td> <td>10</td> <td>BE490435</td>	522	10	BE490435
13	16.4	91.1 <td>546</td> <td>14</td> <td>BG633702</td>	546	14	BG633702
14	16.4	91.1 <td>554</td> <td>28</td> <td>AQ488864</td>	554	28	AQ488864
15	16.4	91.1 <td>598</td> <td>13</td> <td>BQ240732</td>	598	13	BQ240732
16	16.4	91.1 <td>600</td> <td>14</td> <td>CB553914</td>	600	14	CB553914
17	16.4	91.1 <td>623</td> <td>28</td> <td>A2321116</td>	623	28	A2321116
18	16.4	91.1 <td>713</td> <td>9</td> <td>A1296995</td>	713	9	A1296995
19	16.4	91.1 <td>822</td> <td>10</td> <td>BE410139</td>	822	10	BE410139
20	16.4	91.1 <td>847</td> <td>13</td> <td>BUI34166</td>	847	13	BUI34166
21	16.4	91.1 <td>875</td> <td>13</td> <td>BU937820</td>	875	13	BU937820
22	16.4	91.1 <td>898</td> <td>13</td> <td>EX453678</td>	898	13	EX453678
23	16.4	91.1 <td>913</td> <td>29</td> <td>B2578934</td>	913	29	B2578934
24	16.4	91.1 <td>940</td> <td>29</td> <td>B2571854</td>	940	29	B2571854
25	16.4	91.1 <td>1198</td> <td>29</td> <td>B2565652</td>	1198	29	B2565652
26	16.4	91.1 <td>2432</td> <td>11</td> <td>BC028153</td>	2432	11	BC028153
27	16.4	91.1 <td>271</td> <td>28</td> <td>BH641766</td>	271	28	BH641766
28	16.4	91.1 <td>345</td> <td>28</td> <td>B2187917</td>	345	28	B2187917
29	16.4	91.1 <td>417</td> <td>14</td> <td>CB767279</td>	417	14	CB767279
30	16.4	91.1 <td>470</td> <td>28</td> <td>BH320515</td>	470	28	BH320515
31	16.4	91.1 <td>481</td> <td>14</td> <td>CB728483</td>	481	14	CB728483
32	16.4	91.1 <td>500</td> <td>29</td> <td>B2285951</td>	500	29	B2285951
33	16.4	91.1 <td>536</td> <td>28</td> <td>BH328864</td>	536	28	BH328864
34	16.4	91.1 <td>630</td> <td>13</td> <td>BU671252</td>	630	13	BU671252
35	16.4	91.1 <td>647</td> <td>28</td> <td>B2177926</td>	647	28	B2177926
36	16.4	91.1 <td>656</td> <td>10</td> <td>BF289102</td>	656	10	BF289102
37	16.4	91.1 <td>674</td> <td>14</td> <td>CA504539</td>	674	14	CA504539
38	16.4	91.1 <td>711</td> <td>10</td> <td>BG347557</td>	711	10	BG347557
39	16.4	91.1 <td>801</td> <td>29</td> <td>B2283539</td>	801	29	B2283539
40	16.4	91.1 <td>808</td> <td>13</td> <td>BU255259</td>	808	13	BU255259
41	16.4	91.1 <td>817</td> <td>14</td> <td>CD458065</td>	817	14	CD458065
42	16.4	91.1 <td>1064</td> <td>10</td> <td>BE540693</td>	1064	10	BE540693
43	16.4	91.1 <td>212</td> <td>10</td> <td>BF804235</td>	212	10	BF804235
44	16.4	91.1 <td>285</td> <td>14</td> <td>CD265227</td>	285	14	CD265227
45	16.4	91.1 <td>300</td> <td>9</td> <td>AU098914</td>	300	9	AU098914

ALIGNMENTS

RESULT 1
AM663294/C 195 bp mRNA linear EST 06-APR-2000
bh77a10.y1 NCI CGAP GUI Homo sapiens CDNA clone IMAGE:2968794.5'
similar to TR:015471 015471 MONOCYTE INHIBITORY RECEPTOR PRECURSOR.
// mRNA sequence.

ACCESSION AM663294 GI:7455836
VERSION AM663294
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 195)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).

TITLE

Tumor Gene Index

JOURNAL

COMMENT

Unpublished
Other ESTs: bh77a10.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life
Technologies, Inc. CDNA Library Arrayed by: Chitisa Prange, The
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

Found through the I.M.A.G.E. Consortium/LNL at:
 image.lnl.gov/image/html/resources.shtml
 Seq primer: -40RP from Glibco
 High quality sequence stop: 193.

FEATURES

source

1..195
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2968794"
 /issue_type="2 pooled high-grade transitional cell tumors"
 /lab_host="DH10B"
 /clone_id="NCI CGAP GU1"
 /note="Organ: Genitourinary tract; Vector: pCMV-SPORT6; Site_1: Salt; Site_2: North, Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies."

BASE COUNT

54 a 50 c 62 g 25 t 4 others

ORIGIN

Query Match 91.1%; Score 16.4; DB 9; Length 195;
 Best Local Similarity 94.4%; Pred. No. 2.3e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ACCAGGCGCTGCTCTTG 18
 |||||
 65 ACCAGGCGCTGCTCTTG 68

RESULT 2
 T63918/c
 LOCUS 287 bp mRNA linear EST 17-FEB-1995
 DEFINITION YC08C12.1 Stratiogene lung (#937210) Homo sapiens CDNA clone
 IMAGE:80086 5' similar to SP:RRT2_RAT P06768 RETINOL-BINDING
 PROTEIN II, CELLULAR; mRNA sequence.

ACCESSION T63918
 VERSION T63918.1 GI:667783
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE

AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapelli, B., Chisoco, S., Dietrich, N., Dubuque, T., Favello, A., Gish, M., Hawkins, M., Hutteman, M., Kucaba, T., Lacy, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rikkin, L., Rohlfing, T., Schellendberg, K., Soares, M.B., Tan, F., Thierly-Meg, J., Trevaickis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)

TITLE JOURNAL
 MEDLINE 57044478
 PUBMED 8889549
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: ead@wustl.wustl.edu
 Insert Size: 779
 High quality sequence stops: 271 Source: IMAGE Consortium, LNL. This
 clone is available royalty-free through LNL; contact the IMAGE
 Consortium (info@image.lnl.gov) for further information.
 Insert Length: 779 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 271.
 Location/Qualifiers
 1..287
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="GDB:483703"
 /db_xref="taxon:9606"

FEATURES

source

1..287
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="GDB:483703"
 /db_xref="taxon:9606"

/clone="IMAGE:80086"
 /sex="male"
 /dev stage="72 years"
 /lab_host="SOLR cells (Kanamycin resistant)"
 /clone_id="Stratiogene lung (#937210)"
 /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcorI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT, normal lung. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GATTTCGCGACGACG 3' ~3' adaptor sequence: 5' CTCAGTTTCTTTTCTTTTCTTTT 3' "

BASE COUNT

65 a 78 c 71 g 53 t

ORIGIN

Query Match 91.1%; Score 16.4; DB 14; Length 287;
 Best Local Similarity 94.4%; Pred. No. 2.5e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ACCAGGCGCTGCTCTTG 18
 |||||
 62 ACCAGGCGCTGCTCTTG 45

RESULT 3
 AZ081267/c
 LOCUS 308 bp DNA linear GSS 23-AUG-2000
 DEFINITION SLC2A2 Felis catus fca215 Felis catus genomic, genomic survey
 sequence.

ACCESSION AZ081267
 VERSION AZ081267.1 GI:7405601
 KEYWORDS GSS.
 SOURCE Felis catus (cat)
 ORGANISM Felis catus

REFERENCE Mammalia; Eutheria; Carnivora; Fissipedata; Felidae; Felis.
 1 (bases 1 to 308)
 Murphy, M.D., Sun, S., Chen, Z.-Q., Yun, N., Hirschmann, D., Menotti-Raymond, M. and O'Brien, S.J.
 A radiation hybrid map of the cat genome: implications for comparative mapping
 Genome Res. 10 (5), 691-702 (2000)

TITLE JOURNAL
 MEDLINE 20272149
 PUBMED 10810092
 COMMENT Contact: Murphy MJ
 Laboratory of Genomic Diversity
 National Cancer Institute-FCRDC
 P.O. Box B, Bldg. 560, Rm. 11-10, Frederick, MD 21702, USA
 Tel: 301 846 1239
 Fax: 301 846 6327
 Email: murphyw@mail.ncifcrf.gov
 Email: PCR with specific primers.
 Class: PCR with specific primers.
 Location/Qualifiers
 1..308
 /organism="Felis catus"
 /mol_type="Genomic DNA"
 /strain="Fca215"
 /db_xref="taxon:9685"
 /sex="male"
 /clone_id="Felis catus fca215"

FEATURES

source

1..308
 /organism="Felis catus"
 /mol_type="Genomic DNA"
 /strain="Fca215"
 /db_xref="taxon:9685"
 /sex="male"
 /clone_id="Felis catus fca215"

BASE COUNT

57 a 71 c 75 g 105 t

ORIGIN

Query Match 91.1%; Score 16.4; DB 28; Length 308;
 Best Local Similarity 94.4%; Pred. No. 2.6e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ACCAGGCGCTGCTCTTG 18
 |||||
 95 ACCAGGCGCTGCTCTTG 78

RESULT 4
 BY310214
 LOCUS 356 bp mRNA linear EST 11-DEC-2002

DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
BY102124 RIKEN full-length enriched, stroma cell Mus musculus cDNA clone. l320013G01 5', mRNA sequence.	BY102124	BY102124.1	GI:26500551	EST.	Mus musculus (house mouse)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (baae 1 to 356)	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishida, I., Otsu, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gajdardzic, B., Baldarelli, R., Hill, D.P., Bull, C., Hume, D.A., Quackenbush, U., Schmitt, L.M., Knapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusc, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustlich, S., Hirokawa, N., Jackson, I.J., Javie, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Majumet, D.R., Maltas, R., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavani, W.J., Pereira, G., Peole, G., Petrovsky, N., Pillai, R., Pontus, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.U., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verrardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kitahara, T., Konno, H., Nakamura, M., Sakakura, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, C., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shigenaga, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	22354683	22354683.1					
TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT			
	12466851						
CONTACT: Yoshihide Hayashizaki							
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute							
The Institute of Physical and Chemical Research (RIKEN)							
1-7-22 Tsuboto-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan							
Tel: 81-45-503-9222							
Fax: 81-45-503-9215							
Email: genome-res@gsc.riken.go.jp							
URL: http://genome.gsc.riken.go.jp/							
Alkawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, C., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Sakai, K., Murata, M., Nakamura, M., Nomura, K., Nunazaki, R., Ohno, M., Sasaki, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Yamashita, M. and Hayashizaki, Y. Direct							
Submission							
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. genome. 12, 673-677 (2001)							
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)							
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)							
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)							
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.							
Tissues were provided by Takashi Ishikawa (Department of Surgery							

FEATURES	2 Yokohama City University y-9 Fukura, Kanazawa-hu, Yokohama 236-0004 Japan) whose assistance we gratefully acknowledge. Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.					
SOURCE	Location/Qualifiers 1. .356 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="1320013G01" /cell_type="stroma cell" /clone_lib="RIKEN full-length enriched, stroma cell"					
BASE COUNT	58 a 108 c 115 g 75 t					
ORIGIN						
Query Match	91.1%; Score 16.4; DB 13; Length 356;					
Best Local Similarity	94.4%; Prod. No. 2.7e+03;					
Matches	17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
Db	127 AGCAGCGCCGCTGTCTTG 144					
Oy	1 AGCAGCGCCGCTGTCTTG 18 					
RESULT 5	BY664919 364 bp mRNA linear EST 16-DEC-2002					
LOCUS	BY664919 RIKEN full-length enriched, 14.5 days embryo RP+/+					
DEFINITION	Rathke's pouches Mus musculus cDNA clone K720029107.3, mRNA sequence.					
ACCESSION	BY664919					
VERSION	BY664919.1 GI:27039049					
KEYWORDS	EST.					
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 364) Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Otsato, N., Saito, R., Suuki, H., Yamana, I., Kitayama, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., But, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Brdicz, D., Busic, V., Chochoja, C., Corbett, L. E., Cousins, S., Dalla, E., Dirigant, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierecki, R. M., King, B. L., Kongaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Malcais, L., Marchionni, L., McKenzie, L., Miksi, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, U. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shmada, K., Sultana, R., Takemada, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlesstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Komno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Saeki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.					
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)					
JOURNAL	22354683					
MEDLINE	12466851					
PUBMED						
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute					

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-reseqsec.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
M., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN Integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Michelle Brinkmeier and Sally Camper (Dept.
Human Genetics University of Michigan Medical School 4301
MSRB 3 1500 W. Medical Center Dr. Ann Arbor, MI 48109-0638 USA)
whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES

source

Location/Qualifiers
1.364
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="K720029A07"
/tissue_type="Rachke's pouches"
/dev_stage="14.5 days embryo RP+/+"
/clone_lib="RIKEN full-length enriched, 14.5 days embryo
RP+/+ Rachke's pouches"
BASE COUNT 108 a 79 c 75 g 102 t

Query Match 91.1%; Score 16.4; DB 13; Length 364;
Best Local Similarity 94.4%; Pred. No. 2.7e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OR 1 AGCAGCGCGTCTGCTTG 18
|||||
Db 102 AGCAGCGCTGCTGCTTG 119

RESULT 6 366 bp mRNA linear EST 06-DEC-2002
LOCUS BY064920
DEFINITION musculus cDNA clone 1920038D20 5', mRNA sequence.
ACCESSION BY064920
VERSION BY064920.1 GI:26169005
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
1 (bases 1 to 366)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, U., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,

TITLE
JOURNAL
MEDLINE
PUBMED
12454681
22354683

COMMENT

Quackenbush, J., Schriml, L. M., Kanapin, A., Matzuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Brad, D., Brusic, V., Chotila, C., Corbett,
L. E., Cousens, S., Dalla, E., Dragani, J. A., Fletcher, C. F., Forrest,
A., Frazer, K. S., Gasterland, T., Gathbold, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirozane, N., Jackson, I. J.,
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedierski, R. N.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Paven, W. J., Pettes, G.,
Peele, G., Petrovsky, N., Pillai, R., Pontus, J. U., Qi, D.,
Ramchandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sanderlin, A., Schneider, C., Sempke, C. A., Seton,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, L., Walestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wyshaw-Boris, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kohikawa, T., Konno, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-reseqsec.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
M., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN Integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES

source

Location/Qualifiers
1.366
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1920038D20"
/tissue_type="heart"
/dev_stage="17 days embryo"
/clone_lib="RIKEN full-length enriched, 17 days embryo
heart"
BASE COUNT 84 a 103 c 91 g 87 t 1 others

ORIGIN

Query Match 91.1%; Score 16.4; DB 13; Length 366;
 Best Local Similarity 94.4%; Pred. No. 2.7e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 AGCAGCGCGCTGCTCTTG 18
 |||||
 25 AGCAGCGCGCGCTGCTTG 42

RESULT 7

BY661558

LOCUS

BY661558 RIKEN full-length enriched, 14.5 days embryo RP/+
 Rathe's pouches Mus musculus cDNA clone K720011A10 3', mRNA
 sequence.

ACCESSION

BY661558.1 GI:27031604

VERSION

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 373)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nishida, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
 Gojobori, T., Baldarelli, R., Hill, D.P., Bull, C., Hume, D.A.,
 Quackenbush, J., Schmitt, L.M., Kanapin, A., Matsuda, H., Batalov, S.,
 Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochia, C., Corbani,
 L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,
 A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
 Gough, J., Grimond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
 Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,
 King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,
 P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,
 Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring,
 B.Z., Ringwald, M., Sander, A., Schneider, C., Semp, C.A., Setou,
 M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale,
 R.D., Tomita, M., Verardo, R., Wagner, L., Walsby, C., Wang, Y.,
 Watanabe, Y., Wells, C., Wilming, L.G., Wymshaw-Boris, A., Yanagisawa,
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
 Carninci, P., Hayatsu, N., Hirozumi, K., Ishikawa, T., Kono, H., Nakamura,
 M., Sakazume, N., Sato, K., Shiraki, T., Maki, K., Kanai, J., Aizawa, K.,
 Atakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
 Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
 K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
 E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Yoshinide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suicho-cho, Tsunuma-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@ac.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/

Alizawa, K., Akimura, T., Arikawa, T., Carninci, P., Fukuda, S., Hirozumi,
 T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komori, H., Miyazaki, A.,
 Murakami, M., Nakamura, M., Nomura, K., Nomura, K., Ohno, M., Sakai, K.,
 Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
 M., Waki, K., Watanabe, A., Watanabe, M. and Hayashizaki, Y. Direct
 Submission

Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to

FEATURES

source

1..373
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="K720011A10"
 /tissue_type="Rathe's pouches"
 /dev_stage="14.5 days embryo RP/+"
 /clone_id="RIKEN full-length enriched, 14.5 days embryo
 RP/+ Rathe's pouches"
 RP/+ Rathe's pouches"
 110 a 79 c 79 g 105 t

BASE COUNT

110 a 79 c 79 g 105 t

ORIGIN

Query Match 91.1%; Score 16.4; DB 13; Length 373;
 Best Local Similarity 94.4%; Pred. No. 2.7e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 AGCAGCGCGCTGCTCTTG 18
 |||||
 110 AGCAGCGCGCTGCTCTTG 127

LOCUS

BY666019 388 bp mRNA linear EST 06-DEC-2002

DEFINITION

BY666019 RIKEN full-length enriched, 17 days embryo heart Mus

ACCESSION

BY666019 musculus cDNA clone 192004H19 5', mRNA sequence.

VERSION

BY666019.1 GI:26169868

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

1 (bases 1 to 388)

Nishida, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
 Gojobori, T., Baldarelli, R., Hill, D.P., Bull, C., Hume, D.A.,
 Quackenbush, J., Schmitt, L.M., Kanapin, A., Matsuda, H., Batalov, S.,
 Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochia, C., Corbani,
 L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,
 A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
 Gough, J., Grimond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
 Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,
 King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,
 P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,
 Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring,
 B.Z., Ringwald, M., Sander, A., Schneider, C., Semp, C.A., Setou,
 M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale,
 R.D., Tomita, M., Verardo, R., Wagner, L., Walsby, C., Wang, Y.,
 Watanabe, Y., Wells, C., Wilming, L.G., Wymshaw-Boris, A., Yanagisawa,
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
 Carninci, P., Hayatsu, N., Hirozumi, K., Ishikawa, T., Kono, H., Nakamura,
 M., Sakazume, N., Sato, K., Shiraki, T., Maki, K., Kanai, J., Aizawa, K.,
 Atakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
 Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
 K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
 E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE
JOURNAL
MEDLINE
PUBMED
22354683
12466851

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komori, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct
Subcloning
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multichannel sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES

source
1. .388
/organism="Mus musculus"
/mol_type="mRNA"
/accession="U57BL/6J"
/db_xref="taxon:10090"
/clone="1920046419"
/cistype="heart"
/dev_stage="17 days embryo"
/clone_lib="RIKEN full-length enriched, 17 days embryo
heart"

BASE COUNT
66 a 140 c 91 g 88 t 3 others

ORIGIN

Query Match 91.1% Score 16.4; DB 13; Length 388;
Best Local Similarity 94.4%; Pred. No. 2.7e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGCTTG 18
|||||
Db 25 AGCAGCGCGCTGCTTG 42

RESULT 9 AV661443 391 bp mRNA linear EST 16-JAN-2002
LOCUS AV661443 GUC Homo sapiens cDNA clone GUCGSE03 3', mRNA sequence.
DEFINITION
ACCESSION AV661443

VERSION
AV661443.1 GI:9682457
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 391)
Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.,
Xiao, H., Gu, J., Liu, F., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,
Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,
Hu, G., Gu, J., Chen, Z., and Han, Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (126), 15089-15094 (2001)

TITLE
JOURNAL
MEDLINE
PUBMED
11752456
21625106

COMMENT
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

source
1. .391
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GUCGSE03"
/cistype="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT
89 a 107 c 118 g 77 t

ORIGIN

Query Match 91.1% Score 16.4; DB 9; Length 391;
Best Local Similarity 94.4%; Pred. No. 2.7e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGCTTG 18
|||||
Db 297 AGCAGCGCGCTGCTTG 280

RESULT 10 BY223280 424 bp mRNA linear EST 10-DEC-2002
LOCUS BY223280
DEFINITION BY223280 RIKEN full-length enriched, activated spleen Mus musculus
cDNA clone F830302M06 5', mRNA sequence.
ACCESSION BY223280
VERSION BY223280.1 GI:26404383
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 424)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bone, H., Kondo, S.,
Nikaido, I., Oatso, N., Saito, R., Suzuki, H., Yamnaka, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C.,
Gojobori, T., Balderelli, R., Hill, D.P., Bult, C., Hume, D.A.,
Quackenbush, J., Schriml, L.M., Kanapin, A., Malenda, B., Batelov, S.,
Beisel, K.W., Blake, J.A., Bradt, D., Brusa, V., Chothia, C., Corbett,
L.E., Cousin, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,
A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Guernicich, S., Hirokawa, N., Jackson, I.J.,
Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,
King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons

P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontus, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, J. U., Reid, J., Ring, B. Z., Ringwald, M., Sanderlin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K., Sultana, R., Takenaki, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kikukawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Akawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasuniishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

12466851

22354683

JOURNAL MEDLINE PUBMED

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-research.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Aikawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, N., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multipipillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

Location/Qualifiers

1. .424

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NOD"

/db_xref="taxon:10090"

/clone="F830302M06"

/tissue_type="activated spleen"

/clone_lib="RIKEN full-length enriched, activated spleen"

BASE COUNT 80 a 127 c 131 g 86 t

ORIGIN

Query Match 91.1%; Score 16.4; DB 13; Length 424;
Best Local Similarity 94.4%; Pred. No. 2.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGCTCTG 18
|||||
Db 139 AGCAGCGCGCTGCTCTG 156
|||||

RESULT 11

LOCUS B0317016 442 bp mRNA linear EST 17-MAY-2002

DEFINITION IL3-CT0214-150200-076-F11 CT0214 Homo sapiens cDNA, mRNA sequence.

ACCESSION B0317016

VERSION B0317016.1 GI:20922785

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 442)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. P., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H., Brunstein, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

10737800

20202663

JOURNAL MEDLINE PUBMED

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL3&r2=IL3-CT0214-150200-076-F11&r3=2000-02-15&t4=1>)
Seq primer: puc 18 forward.

FEATURES

source

1. .442

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="CT0214"

/note="Organ: colon; Vector: puc18; Site 1: Sma1; Site 2: Sma1; A mini-library was made by cloning products derived from ONESTIES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 104 a 114 c 120 g 104 t

ORIGIN

Query Match 91.1%; Score 16.4; DB 13; Length 442;
Best Local Similarity 94.4%; Pred. No. 2.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGCTCTG 18
|||||
Db 89 AGCAGCGCGCTGCTCTG 106
|||||

RESULT 12

LOCUS BR490435 522 bp mRNA linear EST 31-JUL-2000

DEFINITION WHE0367_D11.G2125 wheat cold-stressed seedling cDNA library

ACCESSION BR490435

Triticum aestivum cDNA clone wheat WHE0367_D11.G21, mRNA sequence.

VERSION BE490435.1 GI:9610064
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Triticum.
 1 (bases 1 to 522)
 REFERENCE Anderson, O.D., Choe, S., Choi, D.W., Close, T.J., Fenton, R.D., Han,
 P.S., Hala, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,
 Seaton, C.L. and Tong, J.C.
 The structure and function of the expressed portion of the wheat
 genome - Cold-stressed seedling cDNA library
 Unpublished
 JOURNAL Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105595773
 Fax: 5105595818
 Email: oanderson@wr.usda.gov
 Sequence have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seq primer: StrataGene SK primer.
 FEATURES
 source
 1. 522
 Location/Qualifiers
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultiar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone="WHE0367.D11.G21"
 /isave_type="Seedling"
 /dev_stage="Five-day old seedling"
 /lab_host="E. coli SOLR"
 /clone_1lb="Wheat cold-stressed seedling cDNA library"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
 Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized
 , germinated and grown aseptically in the dark at room
 temperature on filter paper with water, mycstatin and
 cefotaxime in covered crystallization dishes. Five-day
 old seedlings were transferred to 5 C cold room and kept
 for 48 hr. The tissue, total RNA, and poly(A) RNA were
 prepared, a cDNA library was made, and the cDNA clones
 were in vivo excised to give phagescript phagemids in the
 T7 Close lab (Choi, Close, Fenton) at the University of
 California, Riverside. Plasmid DNA preparations and DNA
 sequencing were performed in the OD Anderson lab (all
 other authors)."
 BASE COUNT 150 a 103 c 140 g 129 t
 ORIGIN
 Query Match 91.1%; Score 16.4; DB 10; Length 522;
 Best Local Similarity 94.4%; Pred. No. 2,9e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 AGCAGGCCGCTGCTTG 18
 |||||
 Db 296 AGCAGGCCGCTGCTTG 313
 RESULT 13
 BG633702/c 546 bp mRNA linear EST 23-APR-2001
 AT29407 Splice AT Drosophila melanogaster adult testes pOTB7
 Drosophila melanogaster cDNA clone AT29407 5 similar to CG8214:
 FBan0008214 located on: 2R 52C2-52C2; 04/09/2001, mRNA sequence.
 ACCESSION BG633702
 VERSION BG633702.1 GI:13761239
 KEYWORDS EST.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

REFERENCE Ephydriidae; Drosophilidae; Drosophila.
 1 (bases 1 to 546)
 AUTHORS Stapleton, M., Brokstein, P., Hong, L., Asparyani, A., Baxter, E., Berman,
 B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V., Farfan,
 D., Frise, E., George, R., Gonzalez, M., Guarin, R., Harris, N., Li, P.,
 Liao, G., Miranda, A., Mista, S., Mungall, C.J., Nunoo, J., Pardo, J.,
 Paragás, V., Park, S., Phuanavong, S., Wan, K., Yu, C., Lewis, S.E.,
 Celniker, S., and Rubin, G.M.
 BDGP/HMT AT Drosophila EST Project
 Unpublished
 JOURNAL Contact: Stapleton, M.
 BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
 hit genomic AE003809; arm:2R [10637159,10898136]
 estimated-cyto:52C1-52D9; 04/09/2001
 Plate: AT.294 row: A column: 7
 High quality sequence sccp: 472.
 FEATURES
 source
 1. 546
 Location/Qualifiers
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="AT29407"
 /sex="male"
 /dev_stage="0-3 day old Ore-R males"
 /lab_host="Plates AT.10-AT.120: DHS-alpha. Plates
 AT.121-AT.319: DHS-alpha Tona"
 /clone_1lb="AT Drosophila melanogaster adult testes pOTB7"
 /note="Organ: ADULT testes; Vector: pOTB7; Site_1: EcoRI;
 Site_2: XhoI; The mRNA for the testis library was made
 from testes and seminal vesicles hand dissected from 0-3
 day old Ore-R males. RNA kindly provided by the lab of
 Margaret Fuller. Sized fractionated cDNAs were directly
 ligated into pOTB7. Plasmid cDNA library."
 BASE COUNT 156 a 138 c 153 g 99 t
 ORIGIN
 Query Match 91.1%; Score 16.4; DB 10; Length 546;
 Best Local Similarity 94.4%; Pred. No. 2,9e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 AGCAGGCCGCTGCTTG 18
 |||||
 Db 336 AGCTGGCCGCTGCTTG 319
 RESULT 14
 AQ48864/c 554 bp DNA linear GSS 24-APR-1999
 LOCUS RPCI-11.266G3.TV RPCI-11 Homo sapiens genomic clone RPCI-11-266G3,
 DEFINITION genomic survey sequence.
 ACCESSION AQ48864
 VERSION AQ48864.1 GI:4674738
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter
 J.C.
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Map Building
 Unpublished
 JOURNAL Other GSSes: RPCI-11-266G3.TV
 CONTACT: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200

Fax: 301 838 0208
 Email: hbe@cligr.org
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@cligr.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Gene Co (<http://www.rgco.com>). BAC end search page: http://www.cligr.org/cdb/human/bac_end_search/bac_end_search.html.
 Seq primer: 17
 Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..554
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="GDB:7601906"
 /db_xref="taxon:9606"
 /clone="RPCI-11-266G3"
 /sex="Male"
 /cell_type="Lymphocytes"
 /clone_lib="RPCI-11"
 /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; RPCI11 Human Male BAC Library"

BASE COUNT

163 a 148 c 84 g 153 t

ORIGIN

Query Match 91.1%; Score 16.4; DB 28; Length 554;
 Best Local Similarity 94.4%; Pred. No. 3e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGTCTCTG 18
 |||||
 Db 416 AGCAGCGCGCTGTCTCTG 399

RESULT 15

BQ240732

LOCUS BQ240732 598 bp mRNA linear EST 03-MAY-2002
 DEFINITION TaB05013B12R TaB05 Triticum aestivum CDNA clone TaB05013B12R, mRNA
 sequence.

ACCESSION

BQ240732

VERSION

BQ240732.1 GI:20436608

KEYWORDS

EST.

SOURCE

Triticum aestivum (bread wheat)

ORGANISM

Triticum aestivum

REFERENCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae

AUTHORS

; Triticaceae; Triticum.

TITLE

1 (bases 1 to 598)

JOURNAL

Cloutier, S.

COMMENT

Wheat functional genomics - Glenlea developing seeds cDNA libraries

UNPUBLISHED

Unpublished

CONTACT

Contact: Dr. Sylvie Cloutier

CEREAL RESEARCH CENTRE

Cereal Research Centre, Agriculture and Agri-food Canada

195 DAFORCE RD

195 DAFORCE RD, WINNIPEG, MB, CANADA R3T 2M9

TEL: (204) 983-2340

Tel: (204) 983-2340

FAX: (204) 983-4604

Fax: (204) 983-4604

EMAIL: scloutier@agr.ca

Email: scloutier@agr.ca

WAS CLONED DIRECTIONALLY

was cloned directionally, not all sequences generated with reverse

PRIMER WERE FROM THE 5' END

primer were from the 5' end (same with forward primer and 3' end).

AVERAGE INSERT SIZE IS >2.0 KB

Average insert size is >2.0 kb

PLATE: 013 ROW: E COLUMN: 12

Plate: 013 row: E column: 12

SEQ PRIMER: M13 REVERSE

Seq primer: M13 Reverse.

/note="Vector: pSPORT-P (Invitrogen Technologies); Site 1:
 NotI; Site 2: MluI; mRNA obtained from wheat seeds of
 cultivar Glenlea 5 days post-anthesis"

BASE COUNT

170 a 121 c 157 g 150 t

ORIGIN

Query Match 91.1%; Score 16.4; DB 13; Length 598;
 Best Local Similarity 94.4%; Pred. No. 3e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCAGCGCGCTGTCTCTG 18
 |||||
 Db 327 AGCAGCGCGCTGTCTCTG 344

Search completed: September 16, 2003, 20:39:53
 Tool time: 915.049 secs

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 17:35:29 ; Search time 438.614 Seconds
(without alignments)
1865.405 Million cell updates/sec

Title: US-09-594-065-1

Perfect score: 20

Sequence: 1 tcaccaccgcgcacaccc 20

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:*

1: gb_ba:*

2: gb_hcg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pac:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sfs:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pac:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sfs:*

28: em_un:*

29: em_vl:*

30: em_hcg_hum:*

31: em_hcg_inv:*

32: em_hcg_other:*

33: em_hcg_mus:*

34: em_hcg_pln:*

35: em_hcg_rtd:*

36: em_hcg_mam:*

37: em_hcg_vrt:*

38: em_sy:*

39: em_hcgo_hum:*

40: em_hcgo_mus:*

41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	798	6 103657	103657 Sequence 5
2	20	100.0	1478	14 S74390	S74390 gBpath/UL27
3	20	100.0	1488	14 S65444	S65444 UL27-glycop
4	20	100.0	1857	14 HS1GBM	K02720 HSV1 (mutan
5	20	100.0	2088	6 E03093	E03093 DNA encodin
6	20	100.0	2586	6 E00358	E00358 DNA encodin
7	20	100.0	2712	6 BD145144	BD145144 Herpes si
8	20	100.0	2712	6 BD165793	BD165793 Vaccine f
9	20	100.0	2712	6 BD165795	BD165795 Vaccine f
10	20	100.0	2715	6 BD145143	BD145143 Herpes si
11	20	100.0	2715	14 AF097023	AF097023 Human her
12	20	100.0	2715	14 AF295528	AF295528 Human her
13	20	100.0	2715	14 AF311740	AF311740 Human her
14	20	100.0	2715	14 HHU48121	U49121 Human herpe
15	20	100.0	2809	14 AF258899	AF258899 Human her
16	20	100.0	2818	14 AF021340	AF021340 Human her
17	20	100.0	2846	14 HHU12172	U12172 Human herpe
18	20	100.0	2846	14 HHU12174	U12174 Human herpe
19	20	100.0	2855	14 HHU12173	U12173 Human herpe
20	20	100.0	2855	14 HHU12175	U12175 Human herpe
21	20	100.0	2925	6 E03115	E03115 DNA encodin
22	20	100.0	3098	6 E00883	E00883 Fragment of
23	20	100.0	3324	14 HS2GB	M14923 HSV2 glycop
24	20	100.0	3461	6 E01195	E01195 DNA sequenc
25	20	100.0	3465	6 E03024	E03024 DNA encodin
26	20	100.0	3465	6 E03092	E03092 DNA encodin
27	20	100.0	3465	6 E03112	E03112 DNA encodin
28	20	100.0	3472	6 BD165792	BD165792 Vaccine f
29	20	100.0	3472	6 BD165794	BD165794 Vaccine f
30	20	100.0	3472	6 I08685	I08685 Sequence 1
31	20	100.0	3472	6 I08686	I08686 Sequence 2
32	20	100.0	3472	14 HS2GB3	M5118 HSV-2 (333)
33	20	100.0	3640	6 I08361	I08361 Sequence 1
34	20	100.0	3643	14 HS1GBP	K03541 HSV-1 (Pat
35	20	100.0	3755	6 E03025	E03025 DNA encodin
36	20	100.0	3755	6 E03113	E03113 DNA encodin
37	20	100.0	3758	14 HS1GB	K01760 HSV1 (KOS)
38	20	100.0	3936	6 E03026	E03026 DNA encodin
39	20	100.0	3996	6 E03114	E03114 DNA encodin
40	20	100.0	3997	6 E00357	E00357 DNA encodin
41	20	100.0	6098	14 HS1GBICP	M14164 HSV1 (strai
42	20	100.0	9756	14 HS1GLYB	M21629 Herpes simp
43	20	100.0	152261	14 HS1GC	X14112 Human herpe
44	20	100.0	154746	14 HSV2RG52	Z86099 Herpes simp
45	18.4	92.0	1521	8 AF148506	AF148506 P15um bat

ALIGNMENTS

RESULT 1

LOCUS 103657

DEFINITION Sequence 5 from Patent US 4642333.

ACCESSION 103657

VERSION 103657.1 GI:268613

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 798)

AUTHORS Person.S.

TITLE Immunologically reactive non-glycosylated amino acid chains of glycoprotein B of herpes virus types 1 and 2

JOURNAL Patent: US 4642333-A 5 10-FEB-1987;

FEATURES 600 Locust La.; State College, PA
Location/Qualifiers
Source 1..798
/organism="unknown"

BASE COUNT 155 a 261 c 254 g 128 t

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 798;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCAACCGTCAGACCTTC 20
|||||
74 TCACCAACCGTCAGACCTTC 93

Db

RESULT 2
LOCUS 574390 1478 bp DNA linear VRL 30-APR-1995
DEFINITION GBpath/UL27-glycoprotein B [herpes simplex virus type 1 HSV-1, ANgpath, Genomic Mutant, 1478 nt].
ACCESSION 574390
VERSION 574390.1 GI:786563
KEYWORDS Human herpesvirus 1
SOURCE Human herpesvirus 1
ORGANISM Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Simplexvirus.

REFERENCE 1 (bases 1 to 1478)
Kovsal,M., Backl,I., Rajcani,J. and Kaerner,H.C.
Replacement of glycoprotein B gene in the herpes simplex virus type 1 strain ANgpath DNA by that originating from nonpathogenic strain KOS reduces the pathogenicity of recombinant virus
Acta Virol. 38 (2), 77-88 (1994)

JOURNAL MEDLINE 95067449
PUBMED 7976866
REMARK Genbank staff at the National Library of Medicine created this entry [NCBI g1bseq 158378] from the original journal article.
This sequence comes from Fig. 8.

FEATURES
source location/Qualifiers
1..1478
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/db_xref="taxon:10298"
1..1478
/partial
/gene="gbpath/UL27"
/note="glycoprotein B"

BASE COUNT 236 a 451 c 528 g 263 t

ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 1478;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCAACCGTCAGACCTTC 20
|||||
456 TCACCAACCGTCAGACCTTC 437

Db

RESULT 3
LOCUS 565444 1488 bp DNA linear VRL 21-APR-2003
DEFINITION UL27-glycoprotein B [herpes simplex virus type 1 HSV-1, ANG, Genomic, 1488 nt].
ACCESSION 565444
VERSION 565444.1 GI:415494
KEYWORDS Human herpesvirus 1
SOURCE Human herpesvirus 1
ORGANISM Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Simplexvirus.

REFERENCE 1 (bases 1 to 1488)

AUTHORS Yuhazs,S.A. and Stevens,J.G.
TITLE Glycoprotein B is a specific determinant of herpes simplex virus type 1 neuroinvasiveness
JOURNAL J Virol. 67 (10), 5948-5954 (1993)
MEDLINE 93381796
PUBMED 8396662
REMARK Genbank staff at the National Library of Medicine created this entry [NCBI g1bseq 137392] from the original journal article.
This sequence comes from Fig. 7.

FEATURES
source location/Qualifiers
1..1488
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/db_xref="taxon:10298"
1..1488
/gene="UL27"
1..1488
/gene="UL27"
/note="This sequence comes from Fig. 7"
/codon_start=1
/product="glycoprotein B"
/protein_id="AAP13961.1"
/db_xref="GI:30027223"

gene
CDS

BASE COUNT 260 a 530 c 459 g 239 t

ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 1488;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCAACCGTCAGACCTTC 20
|||||
1028 TCACCAACCGTCAGACCTTC 1047

Db

RESULT 4
LOCUS HS1GBM 1857 bp DNA linear VRL 02-AUG-1993
DEFINITION HSV1 (mutant strain ts85), glycoprotein B (gB) gene.
ACCESSION K02720
VERSION K02720.1 GI:330087
KEYWORDS glycoprotein.
SOURCE Human herpesvirus 1
ORGANISM Human herpesvirus 1
Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Simplexvirus.

REFERENCE 1 (bases 1 to 1857)
Bzik,D.J., Fox,B.A., Deluca,N.A. and Person,S.
Nucleotide sequence of a region of the herpes simplex virus type 1 gB glycoprotein gene: mutations affecting rate of virus entry and cell fusion
Virology 137 (1), 185-190 (1984)

JOURNAL MEDLINE 84301857
PUBMED 6089415
COMMENT Original source text: HSV1 strain ts85 DNA, clone PTBG-B51.
HSV1 strain ts85 is a mutant of strain HEBM that is temperature dependent for virus growth, production of gB, and fusion of Vero and HEL cells. The syn and ts lesions of ts85 are separable by recombination, the syn defect being located between 0.345 and 0.355 map units, and the ts defect being located between 0.360 and 0.368 map units. The rate of entry determinant of ts85 is genetically separable from both the syn and ts lesions and is located between 0.350 and 0.360 map units, which is entirely within the glycoprotein B gene (0.348 to 0.366 map units).

Strain t8B5 has a total of 12 base substitutions within the coding region for gb when compared to strain KOS. Of these, only two cause amino acid substitutions; a 't' to 'c' transition at bp 695 causes a Val to Ala substitution, and a 'g' to 'a' transition at bp 1610 causes an Arg to His substitution. Because the rate-of-entry determinant has been mapped to between 0.350 and 0.360 map units, the transition at bp 695 must be responsible for the difference in rate-of-entry between t8B5 and KOS.

FEATURES

source

CDS

1..1857
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/db_xref="taxon:10298"
<1..1752
/note="glycoprotein B"
/protein_id="AAA4577.1"
/codon_start=1
/db_xref="GI:330088"
/translation="VDGFYARDUTTKARATAPPTNNLLTPKFTVAMDWPPKPSVCT
MTKQEVDEMISEVGSFSSDAISTFTNTLSEYPLSDICIGKADAMDRI
IFARYNATHIKVGOPOYLIANGFLIAPLPLSLTALRYEHLREOSRKPPTP
PPGASANASVERIKTTSIEBAPLOFTYNNHIOHVNDMLGRVATAMCELOHETLM
NEARLTNNALISATVGRVSRMLGDMVASTCVPVADNVIYONSRISSRPACV
SRPLVSRFEDGGLVEGQLGNNELRLTRDIEFCVGHRRYTFPGGIVYFESYAY
SHQLSRADITTVSTFDLNTLMLDEHEFVPLEVYRHEIKOSGLDYTEVORQNLHD
LRFADIDIVIHADANAMAFAGAFEEGMDGGRAGVKGVIIGCVASVAGSFS
SNPFALAVGLVLGLAALPAFPAYRVMELGSPKALPLTKELKPNPTNDASGEG
EEGGDFDEKALAEAREMTRYMALVSAMETHEKAKKGSRLSLAKYTDWNRKRRNT
NTYVPMNDGDMDDEDL"

BASE COUNT 358 a 624 c 576 g 299 t
ORIGIN 1 bp upstream of SalI site; map coordinate 0.360.

Query Match Best Local Similarity 100.0%; Score 20; DB 14; Length 1857;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCACCACCGTCAGCACCTTC 20
Db 1028 TCACCACCGTCAGCACCTTC 1047

RESULT 5

LOCUS E03093 2088 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding truncated gb of herpes simplex virus type1 (HSV-1).
ACCESSION E03093
VERSION E03093.1 GI:2171311
KEYWORDS JP 1991218397-A/2.
SOURCE Human herpesvirus 1
ORGANISM Human herpesvirus 1
Viruses: dadNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.

REFERENCE 1 (bases 1 to 2088)
AUTHORS Fujisawa,Y., Hinuma,K., Asakawa,N. and Otake,S.
TITLE POLYPEPTIDE AND ITS PRODUCTION
JOURNAL Patent: JP 1991218397-A 2 25-SEP-1991;
TAKEDA CHEM IND LTD
COMMENT OS Herpes simplex virus type1 (HSV-1)
PN JP 1991218397-A/2
PD 25-SEP-1991
PF 21-JUN-1990 JP 1990161448
PR 22-JUN-1989 JP 89P 158238, 30-NOV-1989 JP 89P 308941 PI
FUJISAWA YUKIO, HINUMA KUNIKI, ASAKAWA NAOKI, OTAKE SACHIO PC
C07K13/00 C12N1/19, C12N15/38, C12P21/02//A61K39/245, (C12N1/19, PC
C12N1:865),
PC (C12P21/02, C12R1:865),
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=Miyama,
CC *source: clone=PHS106 delta Tch;
FH Key Location/Qualifiers

PH 1..2085
FT /product='truncated gb of HSV-1' FT 3'UTR
FT 2086..2088
FT mat_peptide 1..2082
FT /product='truncated gb of HSV-1'.
Location/Qualifiers

FEATURES

source

1..2088
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/db_xref="taxon:10298"
BASE COUNT 417 a 751 c 620 g 300 t
ORIGIN

Query Match Best Local Similarity 100.0%; Score 20; DB 6; Length 2088;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCACCACCGTCAGCACCTTC 20
Db 1904 TCACCACCGTCAGCACCTTC 1923

RESULT 6

LOCUS E00358 2586 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding glycoprotein B of herpes simplex virus type 2.
ACCESSION E00358
VERSION E00358.1 GI:2168645
KEYWORDS JP 1985115529-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2586)
AUTHORS Sutanreep,
TITLE NON-GLYCOSTYLATED AMINO ACID CHAIN IMMUNOLOGICALLY REACTIVE TO
GLYCOPROTEIN OF HERPES VIRUS 1 AND 2
JOURNAL Patent: JP 1985115529-A 2 22-JUN-1985;
SUTANREEP PASON
COMMENT OS herpes simplex virus type 2
PN JP 1985115529-A/2
PD 22-JUN-1985
PF 23-JUN-1984 JP 1984122915
PR 23-JUN-1983 US 83 506986, 16-SEP-1983 US 83 532996 PI
SUTANREEP PASON
PC A61K39/245,C07H21/04,C07K15/14,C12N15/00//C12P21/02; CC
Strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HG 52;
CC *source: clone=p52BX1;
CC *source: map_position=0.386-0.348;
FH Key Location/Qualifiers
FT 1..2727
FT /product='glycoprotein B of herpes simplex FT
FT virus type 2'.
Location/Qualifiers

FEATURES

source

1..2586
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:132644"
BASE COUNT 481 a 886 c 841 g 378 t
ORIGIN

Query Match Best Local Similarity 100.0%; Score 20; DB 6; Length 2586;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCACCACCGTCAGCACCTTC 20
Db 1917 TCACCACCGTCAGCACCTTC 1936

```

RESULT 7
BD145144
LOCUS BD145144 2712 bp DNA linear PAT 17-JAN-2003
DEFINITION Herpes simplex virus VP16 vaccine.
ACCESSION BD145144.1 GI:27850902
VERSION JP 2002136297-A/4.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2712)
AUTHORS Burke, R.L. and Sekulovich, R.E.
TITLES Herpes simplex virus VP16 vaccine
JOURNAL Patent: JP 2002136297-A 4 14-MAY-2002;
CHIRON CORP
COMMENT OS HSV gB1
PN JP 2002136297-A/4
PD 14-MAY-2002
PR 10-SEP-2001 JP 2001277435
PR 02-AUG-1990 US 561.528
PI RAE LYN BURKE, ROSE E SEKULOVICH
PC C12N15/09, A61K38/00, A61K39/245, A61P31/12, C07K14/03, C12N15/00,
PC A61K37/02
CC Herpes simplex virus VP16 vaccine
FH Key Location/Qualifiers
FT CDS (1)..(2712).

FEATURES
source
1. .2712
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 513 a 949 c 853 g 397 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 2712;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCAACCTGACGACCTTC 20
|||||
DB 1988 TCACCAACCTGACGACCTTC 2007

RESULT 8
BD165793
LOCUS BD165793 2712 bp DNA linear PAT 17-JAN-2003
DEFINITION Vaccine for use in the therapeutic treatment of HSV.
ACCESSION BD165793
VERSION BD165793.1 GI:27871605
KEYWORDS JP 2002167398-A/2.
SOURCE Herpes simplex virus 7
ORGANISM Herpes simplex virus 7
REFERENCE 1 (bases 1 to 2712)
AUTHORS Burke, R.L., Pachl, C. and Valenzuela, P.D.T.
TITLES Vaccine for use in the therapeutic treatment of HSV
JOURNAL Patent: JP 2002167398-A 2 11-JUN-2002;
CHIRON CORP
COMMENT OS Herpes simplex virus 7
PN JP 2002167398-A/2
PD 11-JUN-2002
PR 12-SEP-2001 JP 2001277340
PR 20-OCT-1986 US 921.213
PI RAE LYN BURKE, CAROL PACHL, PABLO D T VALENZUELA
PC C07K14/035, A61K39/245, A61P31/22, C12N15/09, C12N15/09, C12N15/00, PC
C12N15/00
CC Vaccine for use in the therapeutic treatment of HSV FH Key
FH Key Location/Qualifiers
FT CDS (1)..(2712).

FEATURES
source
1. .2712
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 513 a 949 c 853 g 397 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 2712;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCAACCTGACGACCTTC 20
|||||
DB 1988 TCACCAACCTGACGACCTTC 2007

```

```

source
1. .2712
/organism="Herpes simplex virus 7"
/mol_type="genomic DNA"
/db_xref="taxon:57277"
BASE COUNT 490 a 938 c 884 g 400 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 2712;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCAACCTGACGACCTTC 20
|||||
DB 1982 TCACCAACCTGACGACCTTC 2001

RESULT 9
BD165795
LOCUS BD165795 2712 bp DNA linear PAT 17-JAN-2003
DEFINITION Vaccine for use in the therapeutic treatment of HSV.
ACCESSION BD165795
VERSION BD165795.1 GI:27871607
KEYWORDS JP 2002167398-A/4.
SOURCE Herpes simplex virus 7
ORGANISM Herpes simplex virus 7
REFERENCE 1 (bases 1 to 2712)
AUTHORS Burke, R.L., Pachl, C. and Valenzuela, P.D.T.
TITLES Vaccine for use in the therapeutic treatment of HSV
JOURNAL Patent: JP 2002167398-A 4 11-JUN-2002;
CHIRON CORP
COMMENT OS Herpes simplex virus 7
PN JP 2002167398-A/4
PD 11-JUN-2002
PR 12-SEP-2001 JP 2001277340
PR 20-OCT-1986 US 921.213
PI RAE LYN BURKE, CAROL PACHL, PABLO D T VALENZUELA
PC C07K14/035, A61K39/245, A61P31/22, C12N15/09, C12N15/09, C12N15/00, PC
C12N15/00
CC Vaccine for use in the therapeutic treatment of HSV FH Key
FH Key Location/Qualifiers
FT CDS (1)..(2712).

FEATURES
source
1. .2712
/organism="Herpes simplex virus 7"
/mol_type="genomic DNA"
/db_xref="taxon:57277"
BASE COUNT 514 a 949 c 853 g 396 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 2712;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCAACCTGACGACCTTC 20
|||||
DB 1991 TCACCAACCTGACGACCTTC 2010

RESULT 10
BD145143
LOCUS BD145143 2715 bp DNA linear PAT 17-JAN-2003
DEFINITION Herpes simplex virus VP16 vaccine.
ACCESSION BD145143
VERSION BD145143.1 GI:27850901
KEYWORDS JP 2002136297-A/3.
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2715)
AUTHORS Burke, R.L. and Sekulovich, R.E.
TITLES Herpes simplex virus VP16 vaccine

```


JOURNAL Patent: JP 2002136297-A 3 14-MAY-2002;
CHIRON CORP
COMMENT OS HSV GB2
PN JP 2002136297-A/3
PD 14-MAY-2002
PF 10-SEP-2001 JP 2001274335
PR 02-AUG-1990 US 561,528
PI RAE LYN BURKE, ROSE B SEKULOVICH
PC C12N15/09, A61K39/00, A61K39/245, A61P31/12, C07K14/03, C12N15/00,
PC A61K37/02
CC Herpes simplex virus VP16 vaccine
FH key
FT CDS Location/Qualifiers
1..2715
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32844"
BASE COUNT 491 a 938 c 886 g 400 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 2715;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCACCACCGTCAGACCTTC 20
Db 1982 TCACCACCGTCAGACCTTC 2001
RESULT 11
AF097023 2715 bp DNA linear VRL 07-APR-2000
LOCUS Human herpesvirus 1 strain HSZP glycoprotein B (UL27) gene,
DEFINITION complete cds.
ACCESSION AF097023
VERSION AF097023.1 GI:6165611
KEYWORDS
SOURCE Human herpesvirus 1
ORGANISM Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
1 (bases 1 to 2715)
Kosovsky, J., Vojvodova, A., Oravcova, I., Kudelova, M., Matls, J. and
Rajcani, J.
TITLE Herpes simplex virus 1 (HSV-1) strain HSZP glycoprotein B gene:
JOURNAL comparison of mutations among strains differing in virulence
MEDLINE Virus Genes 20 (1), 27-33 (2000)
PUBMED 1076304
REFERENCE 2 (bases 1 to 2715)
Kosovsky, J., Vojvodova, A., Oravcova, I., Majorova, J., Kudelova, M.
and Rajcani, J.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-1998) Institute of Virology, Slovak Academy of
Sciences, Dubravska cesta 9, Bratislava 842 46, Slovakia
FEATURES
source
1..2715
Location/Qualifiers
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/strain="HSZP"
/db_xref="taxon:10298"
1..2715
/gene="UL27"
1..2715
/gene="UL27"
/gene="UL27"
/function="essential for fusion between virion envelope
and cell membrane"
/note="envelope glycoprotein, gB"
/codon_start=1
/product="glycoprotein B"
/protein_id="AA04615.1"
/db_xref="GI:6165612"

/translation="NRGAPARGCMFVWVALLGLTVASAASSPSTGVAAT
QANAGPATAPAPALGAAPTGDPKPKKKKKPKNPPPPRAGNATVAAGHATLREHLR
DIKASTANFVCPPTGATVVOPEORPCPTREGQNTGIVAPENTAPYEFK
ATMYKDVTVSGVWPGHRYGVSOMGIPEBRAPRPREVIDIKAKVCSTAYVBNL
ETTAAPHRDHETDELKPANAATRTSRGHTTDLKYNPRVDAFHYGTTCYBEV
DARSVPYDFEVLATGDPVMSPFYGRGSHTEHTSYAADFVKDVGFPYADLTTKA
RATAPATRNLLTTPKFTVAMDVPYRPSVCTKQOEVDMLRSYVGSFRPSDAIS
TPTNTNTEYPLSRVLDGICIGKADANDRIFARRYNATHIKVGOPOYYLANGFLI
AYQPLNSTLAEIYREHLREOSRDPNTPPPRASANASVETIKTSSISPARLOF
TNNHQRVNDMLGRVATAKCELOHETLUNBEAKINMNAIASATVERVARNHGD
VWASTCPVPAADNVIVONSRISRPACISRPVVSFRYEDGPIVBSQLEBNELR
LTRDAIERTCTGHRRYFTFGGVYVEBYAASHQSRADITTVSTFIADITLMBE
FVPLEVYTRHEIKDSGLDYTEVQRNQIHDRFADIDTVIHADNAAAMPAGIAFE
GNGDGRAVGVWGIYGVGVSASVGSFMSNPGALAVGLVLAGLAAPFAPRYV
MLOSNPKMAYPLTTKELNPTNPDSAGEGSGDPRDAKLABEMIRVMAVVSAM
EHTERAKKKKGTSALLSAKYTDVMWRKRNTNTYTYVKNKGADBDL"
BASE COUNT 516 a 943 c 856 g 400 t
ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 2715;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCACCACCGTCAGACCTTC 20
Db 1991 TCACCACCGTCAGACCTTC 2010
RESULT 12
AF295528 2715 bp mRNA linear VRL 01-SEP-2001
LOCUS Human herpesvirus 2 glycoprotein B2 (gB2) mRNA, complete cds.
DEFINITION AF295528
ACCESSION AF295528
VERSION AF295528.1 GI:15428575
KEYWORDS
SOURCE Human herpesvirus 2
ORGANISM Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
1 (bases 1 to 2715)
Lee, H.H.
TITLE Sequencing and high level expression of the HSV-2 strain G
JOURNAL glycoprotein B gene using a baculovirus HcNPV
REFERENCE 2 (bases 1 to 2715)
Unpublished
AUTHORS Lee, H.H.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-2000) Department of Biology, Konkuk University, 1
Hwayang-dong, Kwangjin-gu, Seoul 143-701, Korea
FEATURES
source
1..2715
Location/Qualifiers
/organism="Human herpesvirus 2"
/mol_type="mRNA"
/strain="G"
/db_xref="taxon:10310"
1..2715
/gene="gB2"
1..2715
/gene="gB2"
/gene="gB2"
/translation="NRGGGLTCAIVGALVAASAPAPAPAPASGCVATVAANG
GPASRPVPSPATTKARKRKPKPEPEATPPPDANATVAAGHATLRAHLEIKVE
NADAPVPCPEPTGATVVOPEORPCPTREGQNTGIVAPENTAPYEFKATMY
KDVTVSGVWPGHRYGVSOMGIPEBRAPRPREVIDIKAKVCSTAYVBNL
ETTAAPHRDHETDELKPANAATRTSRGHTTDLKYNPRVDAFHYGTTCYBEV
DARSVPYDFEVLATGDPVMSPFYGRGSHTEHTSYAADFVKDVGFPYADLTTKA
RATAPATRNLLTTPKFTVAMDVPYRPSVCTKQOEVDMLRSYVGSFRPSDAIS
TPTNTNTEYPLSRVLDGICIGKADANDRIFARRYNATHIKVGOPOYYLANGFLI
AYQPLNSTLAEIYREHLREOSRDPNTPPPRASANASVETIKTSSISPARLOF
TNNHQRVNDMLGRVATAKCELOHETLUNBEAKINMNAIASATVERVARNHGD
VWASTCPVPAADNVIVONSRISRPACISRPVVSFRYEDGPIVBSQLEBNELR
LTRDAIERTCTGHRRYFTFGGVYVEBYAASHQSRADITTVSTFIADITLMBE
FVPLEVYTRHEIKDSGLDYTEVQRNQIHDRFADIDTVIHADNAAAMPAGIAFE
GNGDGRAVGVWGIYGVGVSASVGSFMSNPGALAVGLVLAGLAAPFAPRYV
MLOSNPKMAYPLTTKELNPTNPDSAGEGSGDPRDAKLABEMIRVMAVVSAM
EHTERAKKKKGTSALLSAKYTDVMWRKRNTNTYTYVKNKGADBDL"

492 A
 939 C 886 G 398 T
 ERTHKAKKTSLLSSKVTNNLRKKNKARYSPLNHDEAGEDEL"

[illegible]

RESULT	13
AF311740	
LOCUS	AF311740 2715 bp DNA linear VRL-24-GAN-2001
DEFINITION	Human herpesvirus 1 strain KOS glycoprotein B gene, complete cds.
ACCESSION	AF311740
VERSION	AF311740.1 GI:11321322
KEYWORDS	.
SOURCE	Human herpesvirus 1
ORGANISM	Human herpesvirus 1

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 2715)	Pertel, P.E., Fridberg, A., Parish, M.L. and Spear, P.C.	Cell fusion induced by herpes simplex virus glycoproteins gB, gD, gC

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

Virology 279 (1), 313-324 (2001)
21066717
11145912
2 (bases 1 to 2715)
Petrel, P.E., Fridberg, A., Parish, M.L. and Spear, P.G.
Direct Submission
Submitted (03-OCT-2000) Microbiology-Immunology, Northwestern
University Medical School, 320 E. Superior Ave.,
Chicago, IL 60611, USA

FEATURES	Location/Qualifiers
source	1. .2715

CDS 1.2715

BASE COUNT	521 a	948 c	849 g	397 t
------------	-------	-------	-------	-------

Query Match	100.0%	Score 20;	DB 14;	Length 2715;
Best Local Similarity	100.0%	Pred. No. 3.2e+02;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TCACACGCTCAGCAGCTTC	20	
Db	1991	TCACACGCTCAGCAGCTTC	2010	

RESULT	14
HU049121	
LOCUS	HU049121
DEFINITION	Human herpesvirus 1 strain ANG path. glycoprotein B (gB) gene,
ACCESSION	complete cds.
VERSION	U49121
FEATURES	U49121.1 GI:1216049

ORGANISM	Human herpesvirus 1 Virusae; dsDNA virusae, no RNA stage; Herpesviridae; Alphaherpesvirinae: Simplexvirus. 1 (bases 1 to 2715)
REFERENCE	Holland T. C. and Scharfch-Longroodi, A. Direct Submission
AUTHORS	Submitted (13-FEB-1996) Dept. of Immunology and Microbiology, Wayne State University, 540 E. Canfield Ave., Detroit, MI 48201, USA
JOURNAL	Location/Qualifiers
FEATURES	

```

/mol_type="genomic DNA"
/strain="ANG path"
/db_xref="taxon:10298"
1..2715
gene

```

```

/notes="HSV-1 UL27 open reading frame"
1..2715
CDS

```

```

CDS
    /note="HSV-1 UL27 open reading frame"
    1..2715
    /gene="gB"
    /note="envelope glycoprotein"
    /codon_start=1
    /product="glycoprotein B"
    /protein_id="AA01805.1"
    /db_xref="GI:1218050"
    /translation="WRGAPRGRGRMYVWALLGLTGVLVSAABSSPGRGVAAAT

```

QAAAGCPATPAPAPGAPPTGDPKPKKNKKKKPKPKBRPAGCNATTAAGHATLREHBR
QAAAGCPATPAPAPGAPPTGDPKPKKNKKKKPKPKBRPAGCNATTAAGHATLREHBR
DIAKENTDANFVCPPTGATVQFQSPRRCPTRPEGONYEGIAVFKENIAPYKPK
DIAKENTDANFVCPPTGATVQFQSPRRCPTRPEGONYEGIAVFKENIAPYKPK

TTFTTNLT EYPLSRVDIGDCIGKDARDAMDRI FARRYNATHIKVGOPYYLANGGFLI

RESULT 15

DEFINITION Human herpesvirus 1 glycoprotein B (UL27) gene, complete cds.
ACCESSION AF259899
VERSION GI:7839508
KEYWORDS
SOURCE
ORGANISM Human herpesvirus 1
Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Simplexvirus.
REFERENCE 1 (bases 1 to 2809)
AUTHORS Ling,J.-Y., Chen,T.-M. and Stroop,W.G.
TITLE A viral genetic element involved in a rabbit model of herpes simplex virus-induced epileptiform seizures
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2809)
AUTHORS Ling,J.-Y., Chen,T.-M. and Stroop,W.G.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2000) Ophthalmology, Mail Slot 523, University of Arkansas for Medical Sciences, 4301 West Markham, Little Rock, AR 72205, USA

FEATURES
source location/Qualifiers
1..2809
/organism="Human herpesvirus 1"
/mol_type="genomic DNA"
/strain="+GC"
/db_xref="taxon:10298"
38..2752
/gene="UL27"
38..2752
/gene="UL27"
/codon_start=1
/product="glycoprotein B"
/protein_id="AAF70301.1"
/db_xref="GI:7839509"
/translation="MRQGAARGCRFWVVALIGLIVLVASAAAPSSRGTPGVAAT
QAANGAPATPAPGAPAPGPTGATVQEBQPRCPTRBQNTTEGLAVKENTAPYFK
DIAKCTDANFYVCPPTGATVQEBQPRCPTRBQNTTEGLAVKENTAPYFK
ATMYKDVVSVQVWFGHRSQFMGI FEDRAVPFSEVIDKINAKGCRSTACTYRNL
ETTAHRDDEHETDMLKPNANATRTSRGWHITDLKYNPSREAFHFGYTTVNCIIEV
DARSVPYNEFVLTADFPYMSPPFYGRSGSHTEHTVSAADRKFQVDGFYARDLTTKA
RATAPTTNLTFTPKFTVAMDVMPKPSVCWTMKQOEVDMLRSEYGSFRSSDAIS
TFTTNLTREYPSRYDLGDCIGKDARDNDRI FARRYNATHIKVGQPOYYLANGSPLI
AYQPLISNTLARIYREHIREQSRKPNFTPPPGASANAVERIKTTSIEFARLQF
TYNHLQRHVNDMLGRVALMCELOHNEHETLWNEARKLNPNALASVTVGRVSAEMIGD
VMAVSTCVPAADNVIVQNSMRISSRPGACSRPLVSFRYEDQGPLVEGQLGENNELR
LTRDAIEPCTVGHRRYFTFGGVYFEEYAYSHOLSRADITTVSTPIDNTIMLEDEH
FVPLEVYRHEIKDSGLDYTEVORRNOLDLRFADIDVTIHDANNAAMFAGIGAPE
GNGDGRVAVGKVMGIVGVSAVSQSSFSMSRPFGALAVGLVLTAGLAAAFARRIV
MKLQSNPMKALYPLTTKELKNPTNPDSGSGEGDGFDAKLAAREMLRYMALVSAM
BETTERAKKKGTSALISAKVTVMVRKRNTVYTVQVPPNKSDADDEDL"

BASE COUNT 536 a 974 c 883 g 416 t
ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 2809;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TCACCAACGCTGACACCTTC 20
|||||
DB 2028 TCACCAACGCTGACACCTTC 2047

BEST AVAILABLE COPY

DR WPI; 1991-328397/45.
 DR P-PSDB; AAR14666.
 PT HSVB polypeptide(s) obt'd. by recombinant DNA techniques -
 PT useful as vaccines against HSV and in diagnosis, can be produced
 PT cheaply and safely.
 XX
 XX
 PS Claim 3; Fig 7; 24pp; Japanese.
 CC The sequence encodes a truncated form of the HSVB polypeptide.
 CC The recombinant protein can be used to prepare vaccines for
 CC prophylaxis of HSV infection and for use in diagnostic kits.
 CC See also AA014478.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 XX
 SQ Sequence 2088 BP; 417 A; 751 C; 620 G; 300 T; 0 other;
 Query Match 100.0%; Score 20; DB 12; Length 2088;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCACCAACGTCAGACCTTC 20
 DB 1904 TCACCAACGTCAGACCTTC 1923
 RESULT 2
 AA62158/C
 ID AA62158 standard; DNA; 2645 BP.
 AC
 XX
 XX
 AC AA62158;
 DT 18-DEC-1998 (first entry)
 DE HSV-2 strain SBS Contig ID 14 DNA sequence.
 XX
 XX
 KW HSV-2 strain SBS; immunological response induction; therapy;
 KW antiviral identification; viral protein inhibitor; ss.
 XX
 OS Herpes simplex virus type 2.
 PH Key Location/Qualifiers
 FT CDS complement (81..2645)
 FT /*tag= a
 FT /product= "ORF#1 protein"
 XX
 XX
 PN MO9820016-A1.
 XX
 PD 14-MAY-1998.
 XX
 PF 31-OCT-1997; 97MO-US20016.
 XX
 PR 09-JUN-1997; 97US-0049018.
 PR 04-NOV-1996; 96US-0030279.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PI Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB;
 PI Esser KM, Leary JI;
 DR WPI; 1998-286847/25.
 DR P-PSDB; AAW72113.
 XX
 PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention
 PT and treatment of infection or inducing immunological response in
 PT mammal
 XX
 PS Claim 1; Page 406-407; 748pp; English.
 CC This sequence represents a Herpes simplex virus type-2 (HSV-2) DNA
 CC sequence of the invention. This sequence was isolated from HSV-2 strain
 CC SBS (deposited as ATCC VR-2546), is designated Contig ID 14, and encodes
 CC a HSV-2 protein. The protein can be used for the treatment or

CC prevention of disease, to induce an immunological response in a mammal or
 CC to identify inhibitors, activators or novel antivirals. Antagonists of
 CC the protein can be used to inhibit a viral polypeptide. The DNA sequence
 CC or a vector containing it can also be used to induce an immunological
 CC response in a mammal.
 XX
 SQ Sequence 2645 BP; 404 A; 837 C; 892 G; 512 T; 0 other;
 Query Match 100.0%; Score 20; DB 19; Length 2645;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCACCAACGTCAGACCTTC 20
 DB 814 TCACCAACGTCAGACCTTC 795
 RESULT 3
 AAN71303
 ID AAN71303 standard; DNA; 2712 BP.
 AC
 XX
 XX
 AC AAN71303;
 DT 25-MAR-2003 (updated)
 DT 30-APR-1991 (first entry)
 DE Herpes Simplex Virus-1 GB from pK8X.
 XX
 XX
 KW Vaccine; prophylaxis; treatment; Herpes Simplex Virus-1;
 KW glycoprotein; gb; ss.
 XX
 XX
 OS Herpes simplex virus type 1 (KOS).
 PN US4642333-A.
 XX
 PD 10-FEB-1987.
 XX
 PF 20-JUN-1984; 84US-0622496.
 XX
 PR 20-JUN-1984; 84US-0622496.
 XX
 XX (PERS/) PERSON S.
 PA
 PI Person S;
 PI
 DR WPI; 1987-056354/08.
 DR P-PSDB; AAP71135.
 XX
 PT Amino acid chain of glyco:protein B of HSV-1 and 2 - prep'd. as
 PT recombinant and used for vaccines for herpes simplex virus types 1
 PT and 2.
 XX
 PS Example; Table 2; 16pp; English.
 CC A pure non-glycosylated amino acid (AA) chain comprising a sequence
 CC corresponding to that occurring in glycoprotein B of HSV-1 or HSV-2
 CC virus which is antigenic to HSV-1 of HSV-2, which contains no more
 CC than 750 AA residues, and which includes AA residues 135-649
 CC inclusive is claimed. It can be used to produce vaccines for
 CC prophylaxis and treatment of HSV-1 and HSV-2.
 CC See AAN71399 for the HSV-2 sequence.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC
 XX
 SQ Sequence 2712 BP; 519 A; 945 C; 849 G; 397 T; 2 other;
 Query Match 100.0%; Score 20; DB 8; Length 2712;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCACCAACGTCAGACCTTC 20
 DB 1988 TCACCAACGTCAGACCTTC 2007

RESULT 4
ID AAN71399 standard; DNA; 2724 BP.
XX
AC AAN71399;
XX
DT 25-MAR-2003 (updated)
DT 30-APR-1991 (first entry)
XX
DE Herpes Simplex Virus-2 GB from p52BXX.
XX
KM Vaccine; prophylaxis; treatment; Herpes Simplex Virus-1;
KM glycoprotein; GB; ss.
XX
OS Herpes simplex virus type 2 HG52.
XX
PN USA642333-A.
XX
PD 10-FEB-1987.
XX
PF 20-JUN-1984; 84US-0622496.
XX
PR 20-JUN-1984; 84US-0622496.
XX
PA (PERS/) PERSON S.
XX
PI Person S;
XX
DR WPI; 1987-056354/08.
DR P-PSDB; AAF71135.
XX
PT Amino acid chain of glyco:protein B of HSV-1 and 2 - prepd. as
PT recombinant and used for vaccines for herpes simplex virus types 1
PT and 2.
XX
PS Example; Table 2; 16pp; English.
XX
CC 'N' indicated nucleotide to be determined (corresp. to residues
CC 17-41 of HSV-1 GB).
CC
CC A pure non-glycosylated amino acid (AA) chain comprising a sequence
CC corresponding to that occurring in glycoprotein B of HSV-1 or HSV-2
CC virus which is antigenic to HSV-1 of HSV-2, which contains no more
CC than 750 AA residues, and which includes AA residues 135-649
CC inclusive is claimed. It can be used to produce vaccines for
CC prophylaxis and treatment of HSV-1 and HSV-2.
CC See AAN71303 for the HSV-1 sequence.
CC (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 2724 BP; 494 A; 911 C; 854 G; 390 T; 75 other;
XX
Query Match 100.0%; Score 20; DB 8; Length 2724;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACCAACCGTCAGCACCTTC 20
Db 1991 TCACCAACCGTCAGCACCTTC 2010
XX
RESULT 5
ID AAN50517 standard; DNA; 2727 BP.
XX
AC AAN50517;
XX
DT 21-NOV-1991 (first entry)
XX
DE Sequence encoding the glycoprotein B (GB) of Herpes simplex virus
DE type 2 (HSV-2).
XX
KM Vaccine; glycoprotein B; ss.
XX

OS Herpes simplex virus type 2.
XX
XX Key Location/Qualifiers
FT CDS 1..2727
FT /*tag= a
XX
XX EPI33063-A.
XX
PD 13-FEB-1985.
XX
PF 22-JUN-1984; 84BP-0401312.
XX
PR 16-SEP-1983; 83US-0532996.
PR 23-JUN-1983; 83US-0506986.
PR 20-JUN-1984; 84US-0622496.
XX
PA (PERS/) PERSON S.
XX
PI Person S;
XX
DR WPI; 1985-039636/07.
XX
PT New amino acid chains of glyco:protein B of herpes simplex virus
PT - prepd. by recombination DNA methods for use in vaccines
XX
PS Disclosure; Table 2, Page 35-35C; 40pp; English.
XX
CC The inventors claim a non-glycosylated amino acid chain comprising a
CC sequence corresponding to that occurring in glycoprotein B of HSV-1
CC or HSV-2. Preferably, when the chain comprises a sequence corresp.
CC to that occurring in glycoprotein B of HSV-2, it contains 4-750 AA
CC residues. The chain esp. has a molecular wt. of 65000 daltons and
CC includes AA residues 165-629 from the HSV-2 glycoprotein B sequence.
XX
SQ Sequence 2727 BP; 493 A; 913 C; 855 G; 391 T; 75 other;
XX
Query Match 100.0%; Score 20; DB 6; Length 2727;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACCAACCGTCAGCACCTTC 20
Db 1994 TCACCAACCGTCAGCACCTTC 2013
XX
RESULT 6
ID AAV62146 standard; DNA; 2841 BP.
XX
AC AAV62146;
XX
DT 07-DEC-1998 (first entry)
XX
DE HSV-2 strain SBS Contig ID 91 DNA sequence.
XX
KM HSV-2 strain SBS; immunological response induction; therapy;
KM antiviral identification; viral protein inhibitor; ss.
XX
OS Herpes simplex virus type 2.
XX
XX Key Location/Qualifiers
FH CDS 358..2745
FT /*tag= a
XX
XX WO9820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US20016.
XX
PR 09-JUN-1997; 97US-0049018.
PR 04-NOV-1996; 96US-0030279.
XX

PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 CC Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB;
 PI Esser KM, Leary JJ;
 XX
 DR WPI; 1998-286847/25.
 DR P-PSDB; AAW72062.
 XX
 PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention
 PT and treatment of infection or inducing immunological response in
 PT mammal
 XX
 PS Claim 1; Page 284-285; 748pp; English.
 XX
 CC This sequence represents a Herpes simplex virus type-2 (HSV-2) DNA
 CC sequence of the invention. This sequence was isolated from HSV-2 strain
 CC SBS (deposited as ATCC VR-2546), is designated Contig ID 91, and encodes
 CC a HSV-2 protein. The proteins can be used for the treatment or
 CC prevention of disease, to induce an immunological response in a mammal or
 CC to identify inhibitors, activators or novel antivirals. Antagonists of
 CC the protein can be used to inhibit a viral polypeptide. The DNA sequence
 CC or a vector containing it can also be used to induce an immunological
 CC response in a mammal.
 XX
 SQ Sequence 2841 BP; 577 A; 921 C; 915 G; 425 T; 3 other;
 XX
 Query Match 100.0%; Score 20; DB 19; Length 2841;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 TCACCACTGTCAGACCTTC 20
 DB 2012 TCACCACTGTCAGACCTTC 2031
 XX
 RESULT 7
 ID AAQ14455
 XX AAQ14455 standard; DNA; 2925 BP.
 XX
 AC AAQ14455;
 XX
 DT 25-MAR-2003 (updated)
 DT 21-JAN-1992 (first entry)
 XX
 DE HSV surface antigens gB/gD fusion protein gene.
 XX
 KM Herpes simplex virus; vaccine; ss.
 XX
 OS Herpes simplex virus.
 XX
 PI Key Location/Qualifiers
 FH CDS 1..2925
 FT /*tag= a
 PT
 XX
 PN JF03220200-A.
 XX
 PD 27-SEP-1991.
 XX
 PF 29-NOV-1990; 90JP-0325474.
 XX
 PR 30-NOV-1989; 89JP-0308942.
 PR 29-NOV-1990; 90JP-0325474.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 DR WPI; 1991-329235/45.
 DR P-PSDB; AAR14680.
 XX
 PT Polypeptide for herpes simplex virus vaccine prodn. - obid. by
 PT culturing transformant of recombinant DNA coding fused protein of
 PT virus surface antigen
 XX
 PS Claim 6; Fig 8; 16pp; Japanese.
 f

XX
 CC The DNA sequence encodes a fusion protein comprising herpes simplex
 CC virus (HSV) surface antigens gB and gD. This fusion protein shows
 CC the same physiological activity as surface antigen prep. from HSV
 CC infected cells. It can be used in a vaccine for the prevention of
 CC HSV infection.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 2925 BP; 579 A; 1061 C; 848 G; 437 T; 0 other;
 XX
 Query Match 100.0%; Score 20; DB 12; Length 2925;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 TCACCACTGTCAGACCTTC 20
 DB 1904 TCACCACTGTCAGACCTTC 1923
 XX
 RESULT 8
 ID AAN60195
 XX AAN60195 standard; DNA; 3098 BP.
 XX
 AC AAN60195;
 XX
 DT 25-MAR-2003 (updated)
 DT 31-JUL-1991 (first entry)
 XX
 DE Herpes simplex virus glycoprotein gB gene.
 XX
 KM HSV; gB glycoprotein; vaccine; ss.
 XX
 OS Herpes simplex virus.
 XX
 PI Key Location/Qualifiers
 FH CDS 257..2956
 FT /*tag= a
 PT
 XX
 PN EP170169-A.
 XX
 PD 05-FEB-1986.
 XX
 PF 19-JUL-1985; 85EP-0109042.
 XX
 PR 20-JUL-1984; 84JP-0151766.
 PR 11-DEC-1984; 84JP-0262465.
 XX
 PA (KAGA) CHEMO SERO THERAPEUTIC RES INST.
 XX
 PI Nozaki C, Makizumi K, Kino Y, Eto, T, Ohtomo N;
 XX
 DR WPI; 1986-036935/06.
 DR P-PSDB; AAP60244.
 XX
 PT Recombinant DNA containing herpes simplex virus gene or fragment -
 PT useful in transformant yeast for prodn. of high purity herpes
 PT simplex virus glycoproteins for use in protective vaccines.
 XX
 PS Disclosure; Fig. 6; 53pp; English.
 XX
 CC The gene encodes a gB glycoprotein which is useful for the
 CC production of vaccines conferring protection against herpes simplex
 CC virus infections. The vaccine is safe in use and does not contain a
 CC carcinogenic gene, as with prior vaccines.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 3098 BP; 586 A; 1070 C; 959 G; 483 T; 0 other;
 XX
 Query Match 100.0%; Score 20; DB 7; Length 3098;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 f

QY 1 TCACGACCGTCAGACCTTC 20
DB 2244 TCACGACCGTCAGACCTTC 2263

RESULT 9

ID AAN70681 standard; DNA; 3461 BP.

AC AAN70681;

DT 25-MAR-2003 (updated)

DT 10-MAR-2003 (updated)

DT 30-APR-1991 (first entry)

DE Recombinant herpes simplex virus gene.

KM Herpes simplex virus; E.coli; plasmid; vector; promoter; antigen;

KM vaccine; glycoprotein; HSV gB; ss.

OS Herpes simplex virus.

FH Key Location/Qualifiers

FT CDS 254..2962

FT /*tag= a

FT /product= HSV_gB

JP62115288-A.

PD 26-MAY-1987.

PF 14-NOV-1985; 85JP-0255971.

PR 14-NOV-1985; 85JP-0255971.

PA (KAGA) KAGAKU OYOBI KESSEI RYOKO.

DR WPI; 1987-183377/26.

DR P-PSDB; AAF70426.

PT Recombinant DNA contg. herpes simplex virus gene - is integrated in

PT E.coli plasmid vector, contg. incorporated promoter region for

PT animal cells, giving antigen.

PS Dielosure; Fig 3(1-3); 13pp; Japanese.

CC The gene encoding HSV gB is cloned into E.coli. To increase the

CC efficiency of expression, the gene encoding the C-terminal region

CC of gB, which is very hydrophobic, is removed and the remaining gene

CC is cloned into an expression vector. The vector is composed of pBR322,

CC SV40 early promoter and the HSV gB gene.

CC The construct is useful for the production of a good antigen for

CC a HSV vaccine.

CC (updated on 10-MAR-2003 to add missing OS field.)

CC (updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 3461 BP; 666 A; 1188 C; 1041 G; 566 T; 0 other;

Query Match 100.0%; Score 20; DB 8; Length 3461;

Best Local Similarity 100.0%; Pred. No. 25;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACGACCGTCAGACCTTC 20

DB 2241 TCACGACCGTCAGACCTTC 2260

RESULT 10

AAO14478

ID AAO14478 standard; DNA; 3465 BP.

AC AAO14478;

DT 25-MAR-2003 (updated)

DT 23-JAN-1992 (first entry)

DE HSVgB gene.

KM Vaccine; antigen; ss.

OS Herpes simplex virus.

FH Key Location/Qualifiers

FT CDS 251..2965

FT /*tag= a

JP03218397-A.

PD 25-SEP-1991.

PF 21-JUN-1990; 90JP-0161448.

PR 30-NOV-1989; 89JP-0308941.

PR 22-JUN-1989; 89JP-0158238.

PR 21-JUN-1990; 90JP-0161448.

PA (TAKA) TAKEDA CHEM IND LTD.

DR WPI; 1991-328397/45.

DR P-PSDB; AAR14665.

PT HSVgB polypeptide(s) obt'd. by recombinant DNA techniques -

PT useful as vaccines against HSV and in diagnosis, can be produced

PT cheaply and safely.

PS Claim 3; Fig 1; 24pp; Japanese.

CC The sequence encodes the HSVgB polypeptide. The recombinant protein

CC or esp. truncated forms of it, can be used to prepare vaccines for

CC prophylaxis of HSV infection and for use in diagnostic kits.

CC See also AAO14479.

CC (updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 3465 BP; 666 A; 1194 C; 1044 G; 561 T; 0 other;

Query Match 100.0%; Score 20; DB 12; Length 3465;

Best Local Similarity 100.0%; Pred. No. 25;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACGACCGTCAGACCTTC 20

DB 2241 TCACGACCGTCAGACCTTC 2260

RESULT 11

AAT93650

ID AAT93650 standard; DNA; 3471 BP.

AC AAT93650;

DT 26-FEB-1998 (first entry)

DE Herpes simplex virus type 1 glycoprotein B encoding DNA.

KM Herpes simplex virus; HSV; type I; type II; glycoprotein B; vaccine;

KM membrane protein; virus-specific glycoprotein;

KM transmembrane anchor region; ss.

OS Herpes simplex virus type 1.

FH Key Location/Qualifiers

FT CAAT_signal

FT /*tag= a

FT /note= "putative"

FT TATA_signal

FT /*tag= b

```

FT      CDS                               /note="putative"
FT      309..3023                         /tag= c
FT      polyA_signal                      /tag= d
FT      3166..3173                         /note="putative"
FT      polyA_signal                      /tag= e
FT      3409..3416                         /note="putative"
FT      /note="putative"
XX      US5648079-A.
XX      15-JUL-1997.
XX      08-DEC-1994; 94US-0351875.
XX      20-OCT-1986; 86US-0921730.
XX      06-APR-1984; 84US-0597784.
XX      20-SEP-1990; 90US-0587179.
XX      21-DEC-1992; 92US-0993415.
XX      18-OCT-1993; 93US-0138717.
XX      (CHIR ) CHIRON CORP.
XX      Burke LR, Pachl C, Valenzuela PD;
XX      WPI; 1997-372032/34.
XX      P-PSDB; AAM34552.
XX      Vaccines against herpes simplex virus infection - containing
XX      recombinant HSV glycoprotein B
XX      Disclosure; Fig 4A-F; 33pp; English.
XX      The present sequence is from the Herpes simplex virus (HSV) type I
XX      strain Patcon and encodes a glycoprotein B. HSV is a double stranded
XX      virus packaged within an icosahedral nucleocapsid enveloped within a
XX      membrane. The membrane includes a number of virus-specific
XX      glycoproteins, with glycoprotein B being one of the most abundant.
XX      Glycoprotein B from both HSV type I and type II are cross reactive. New
XX      vaccines comprising recombinantly produced glycosylated glycoprotein B
XX      that has a deletion of at least part of the transmembrane anchor region,
XX      in combination with a carrier and an adjuvant have been produced. The
XX      vaccines are used for immunising humans against HSV (HSV type 1 or
XX      HSV type 2) by vaccination before or after primary infection with HSV.
XX      Sequence 3471 BP; 639 A; 1201 C; 1069 G; 562 T; 0 other;
XX
XX      Query Match          100.0%; Score 20; DB 18; Length 3471;
XX      Best Local Similarity 100.0%; Pred. No. 25;
XX      Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      1 TCACACCGTCGACGACCTTC 20
XX      2290 TCACACCGTCGACGACCTTC 2309
XX
XX      RESULT 12
XX      ID AAN80907 standard; DNA; 3472 BP.
XX      AAN80907;
XX      25-MAR-2003 (updated)
XX      09-MAR-1992 (first entry)
XX      Sequence of Herpes simplex virus (HSV) glycoprotein B (GB) 1 DNA.
XX      Vaccine; herpes simplex virus; therapy; ss.
XX      Herpes simplex virus.
XX      Key Location/Qualifiers
XX

```

```

FT      CDS                               309..3023
FT      /tag= a
XX      MO8802634-A.
XX      21-APR-1988.
XX      20-OCT-1987; 87WO-US02709.
XX      20-JUL-1987; 87US-0079605.
XX      20-OCT-1986; 86US-0921213.
XX      (CHIR ) CHIRON CORP.
XX      Burke LR, Pachl C, Valenzuela P;
XX      WPI; 1988-119368/17.
XX      P-PSDB; AAP80914.
XX      Vaccine for treatment of herpes simplex virus - contains
XX      recombinant HSV glycoproteins B and D
XX      Example; Fig 4; 71pp; English.
XX      Prep. of recombinant GB and GD is described in WO88504587. The
XX      amino acid sequence and DNA sequence for GB 1 presented in AAP80914
XX      and AAN80907 differ from that originally presented in Table 1 of
XX      International Publication No. WO 85/04587. The DNA sequence in Table
XX      1 contains an error in that an additional nucleotide (G) is listed
XX      at position 607 which resulted in a shift in reading frame relative
XX      to AAN80907 from which this nucleotide has been deleted.
XX      (Updated on 25-MAR-2003 to correct PA field.)
XX
XX      Sequence 3472 BP; 639 A; 1202 C; 1068 G; 563 T; 0 other;
XX
XX      Query Match          100.0%; Score 20; DB 9; Length 3472;
XX      Best Local Similarity 100.0%; Pred. No. 25;
XX      Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      1 TCACACCGTCGACGACCTTC 20
XX      2290 TCACACCGTCGACGACCTTC 2309
XX
XX      RESULT 13
XX      ID AAQ48496 standard; DNA; 3472 BP.
XX      AAQ48496;
XX      25-MAR-2003 (updated)
XX      29-MAR-1994 (first entry)
XX      Glycoprotein B (GB1).
XX      Glycoprotein; GB1; GB2; herpes simplex virus; HSV-1; HSV-2;
XX      flanking; initiation; termination; transcription; translation;
XX      regulatory sequence; ss.
XX      Herpes simplex virus 1 strain Patcon.
XX      Key Location/Qualifiers
XX      CDS 309..3023
XX      /tag= a
XX      US5244792-A.
XX      14-SEP-1993.
XX      20-SEP-1990; 90US-0587179.
XX      06-APR-1984; 84US-0597784.
XX      20-OCT-1986; 86US-0921730.
XX

```

PR 20-SEP-1990; 90US-0587179.
XX
XX (CHIR) CHIRON CORP.
PI
XX Burke RL, Pachl C, Valenzuela PDT;
XX
XX WPI: 1993-302641/38.
DR
XX P-PSDB; AAR411778.
XX
XX DNA construct for expressing HSV glycosylated polypeptide -
PT useful for vaccinating against HSV-1 and -2 infections
XX
XX Disclosure; Fig 4; 33pp; English.
XX
XX DNA constructs for expressing a glycosylated polypeptide in a
CC eukaryotic host cell comprises (1) an oligonucleotide sequence (OS)
CC free of natural flanking sequences, encoding glycoprotein B (GB)
CC of HSV, or its fragments, and (2) 5' initiating and 3' terminating
CC transcriptional and translational regulatory sequences flanking OS,
CC at least one of these sequences not being from HSV. The HSV-1 and
CC HSV-2 sequences are given in AAQ48496 and AAQ48497 respectively.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 3472 BP; 639 A; 1202 C; 1069 G; 562 T; 0 other;
Query Match 100.0%; Score 20; DB 14; Length 3472;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACCAACGCTCAGCACCTTC 20
DB 2290 TCACCAACGCTCAGCACCTTC 2309
RESULT 14
AAQ48497 standard; DNA; 3472 BP.
XX
XX AAQ48497;
XX
XX 25-MAR-2003 (updated)
DT 29-MAR-1994 (first entry)
XX
XX Glycoprotein B (GB2).
XX
XX Glycoprotein; GB1; GB2; herpes simplex virus; HSV-1; HSV-2;
KW flanking; initiation; termination; transcription; translation;
KW regulatory sequence; ss.
XX
XX Herpes simplex virus 2 strain 333.
OS
XX
XX Key Location/Qualifiers
FH CDS 309..3023
FT /*tag= a
XX
XX US5244792-A.
XX
XX 14-SEP-1993.
XX
XX 20-SEP-1990; 90US-0587179.
XX
XX 06-APR-1984; 84US-0597784.
PR 20-OCT-1986; 86US-0921730.
PR 20-SEP-1990; 90US-0587179.
XX
XX (CHIR) CHIRON CORP.
XX
XX Burke RL, Pachl C, Valenzuela PDT;
XX
XX WPI: 1993-302641/38.
DR P-PSDB; AAR411779.
XX
XX DNA construct for expressing HSV glycosylated polypeptide -

PT useful for vaccinating against HSV-1 and -2 infections
XX
XX Disclosure; Fig 4; 33pp; English.
XX
XX DNA constructs for expressing a glycosylated polypeptide in a
CC eukaryotic host cell comprises (1) an oligonucleotide sequence (OS)
CC free of natural flanking sequences, encoding glycoprotein B (GB)
CC of HSV, or its fragments, and (2) 5' initiating and 3' terminating
CC transcriptional and translational regulatory sequences flanking OS,
CC at least one of these sequences not being from HSV. The HSV-1 and
CC HSV-2 sequences are given in AAQ48496 and AAQ48497 respectively.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 3472 BP; 661 A; 1193 C; 1053 G; 565 T; 0 other;
Query Match 100.0%; Score 20; DB 14; Length 3472;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACCAACGCTCAGCACCTTC 20
DB 2299 TCACCAACGCTCAGCACCTTC 2318
RESULT 15
AAT93651
ID AAT93651 standard; DNA; 3473 BP.
XX
XX AAT93651;
XX
XX 26-FEB-1998 (first entry)
XX
XX Herpes simplex virus type 2 glycoprotein B encoding DNA.
XX
XX Herpes simplex virus; HSV; type I; type II; glycoprotein B; vaccine;
KW membrane protein; virus-specific glycoprotein;
KW transmembrane anchor region; ss.
XX
XX Herpes simplex virus type 2.
XX
XX Key Location/Qualifiers
FH CAAAT_signal 55..62
FT /*tag= a
FT /note= "putative"
FT TATA_signal 125..131
FT /*tag= b
FT /note= "putative"
FT CDS 309..3023
FT /*tag= c
FT polyA_signal 2744..2751
FT /*tag= d
FT /note= "putative"
XX
XX US5648079-A.
XX
XX 15-JUL-1997.
XX
XX 08-DEC-1994; 94US-0351875.
XX
XX 20-OCT-1986; 86US-0921730.
PR 06-APR-1984; 84US-0597784.
PR 20-SEP-1990; 90US-0587179.
PR 21-DEC-1992; 92US-0993415.
PR 18-OCT-1993; 93US-0138717.
XX
XX (CHIR) CHIRON CORP.
XX
XX Burke RL, Pachl C, Valenzuela PDT;
XX
XX WPI: 1997-372022/34.
DR P-PSDB; AAW34553.
XX
XX Vaccines against herpes simplex virus infection - containing

PT recombinant HSV glycoprotein B
XX
PS Disclosure; Fig 4A-F; 33pp; English.
XX
CC The present sequence is from the Herpes simplex virus (HSV) type 2
CC strain 333 and encodes a glycoprotein B. HSV is a double stranded
CC virus packaged within an icosahedral nucleocapsid enveloped within a
CC membrane. The membrane includes a number of virus-specific
CC glycoproteins, with glycoprotein B being one of the most abundant.
CC Glycoprotein B from both HSV type I and type II are cross reactive. New
CC vaccines comprising recombinantly produced glycosylated glycoprotein B
CC that has a deletion of at least part of the transmembrane anchor region,
CC in combination with a carrier and an adjuvant have been produced. The
CC vaccines are used for immunising humans against HSV (HSV type 1 or
CC HSV type 2) by vaccination before or after primary infection with HSV.
XX
SQ Sequence 3473 BP; 661 A; 1195 C; 1053 G; 564 T; 0 other;
Query Match 100.0%; Score 20; DB 18; Length 3473;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACGACCGTCAGCACCTTC 20
|||
DB 2299 TCACGACCGTCAGCACCTTC 2318

Search completed: September 16, 2003, 18:36:57
Job time : 102.198 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 16, 2003, 18:24:35 ; Search time 25.5446 Seconds
(without alignments)
345.579 Million cell updates/sec

Title: US-09-594-065-1

Perfect score: 20

Sequence: 1 tcaccaccgtcagcacccttc 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/ECTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	2713	US-08-804-439A-13	Sequence 13, Appl
2	20	100.0	2713	US-08-720-229-13	Sequence 13, Appl
3	20	100.0	3472	5244792-1	Patent No. 5244792
4	20	100.0	3472	5244792-2	Patent No. 5244792
5	18.4	92.0	2088	US-08-164-839-7	Sequence 7, Appl
6	18.4	92.0	2088	US-08-583-799-7	Sequence 7, Appl
7	18.4	92.0	2921	US-08-164-839-9	Sequence 9, Appl
8	18.4	92.0	2921	US-08-583-799-9	Sequence 9, Appl
9	18.4	92.0	34094	US-09-292-034-1	Sequence 1, Appl
10	16.8	84.0	788	US-09-142-623-10	Sequence 10, Appl
11	16.8	84.0	1905	US-09-142-623-2	Sequence 2, Appl
12	16.8	84.0	4403765	US-09-103-840A-2	Sequence 2, Appl
13	16.8	84.0	4411529	US-09-103-840A-1	Sequence 1, Appl
14	16.4	82.0	2508	US-09-252-991A-793	Sequence 793, App
15	16.4	82.0	2508	US-09-252-991A-828	Sequence 828, App
16	15.8	79.0	1586	US-09-673-018-1	Sequence 1, Appl
17	15.8	79.0	2124	US-09-252-991A-12830	Sequence 12830, A
18	15.8	79.0	2454	US-08-890-615-1	Sequence 1, Appl
19	15.8	79.0	2454	US-09-246-290A-1	Sequence 1, Appl
20	15.8	79.0	2475	US-09-045-185-1	Sequence 1, Appl
21	15.8	79.0	2475	US-09-045-185-3	Sequence 3, Appl
22	15.4	77.0	264	US-09-252-991A-12642	Sequence 12642, A
23	15.4	77.0	1116	US-09-252-991A-13319	Sequence 13319, A
24	15.4	77.0	1491	US-08-941-647A-4	Sequence 4, Appl
25	15.4	77.0	1505	US-09-142-481-1	Sequence 1, Appl
26	15.4	77.0	2175	US-08-914-999-3	Sequence 3, Appl
27	15.4	77.0	2235	US-09-252-991A-12994	Sequence 12994, A

28	15.4	77.0	3486	US-09-252-991A-12477	Sequence 12477, A
29	15.4	77.0	25165	US-09-453-702B-39	Sequence 39, Appl
30	15.4	77.0	169998	US-09-676-610B-24	Sequence 24, Appl
31	15.4	77.0	197496	US-09-877-177A-10	Sequence 10, Appl
32	15.4	77.0	4403765	US-09-103-840A-2	Sequence 2, Appl
33	15.4	77.0	4411529	US-09-103-840A-1	Sequence 1, Appl
34	15.2	76.0	36	US-09-689-291A-9	Sequence 9, Appl
35	15.2	76.0	100	US-09-298-886-20	Sequence 20, Appl
36	15.2	76.0	135	US-08-148-708-4	Sequence 4, Appl
37	15.2	76.0	291	US-09-313-294A-7433	Sequence 7433, Ap
38	15.2	76.0	294	US-09-313-294A-6329	Sequence 6329, Ap
39	15.2	76.0	296	US-09-313-294A-6788	Sequence 6788, Ap
40	15.2	76.0	377	US-08-332-766A-1	Sequence 1, Appl
41	15.2	76.0	420	US-09-252-991A-2369	Sequence 2369, Ap
42	15.2	76.0	429	US-09-262-537-22	Sequence 22, Appl
43	15.2	76.0	435	US-09-252-991A-2833	Sequence 2833, Ap
44	15.2	76.0	461	US-09-615-192A-375	Sequence 375, App
45	15.2	76.0	497	US-09-439-554-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-804-439A-13
; Sequence 13, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Martin L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439A
; FILING DATE: February 21, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Hallie, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09176/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; TEXT:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-804-439A-13
Query Match 100.0%; Score 20; DB 3; Length 2713;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 TCACCACCGTCGACGACCTTC 20
Db 1988 TCACCACCGTCGACGACCTTC 2007

RESULT 2
US-08-720-229-13

Sequence 13, Application US/08720229
Patent No. 6022542

GENERAL INFORMATION:

APPLICANT: Rose, Timothy M.

APPLICANT: Boesch, Marnix L.

APPLICANT: Strand, Kurt

TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHV/KSHV

NUMBER OF SEQUENCES: 100

CORRESPONDENCE ADDRESS:

ADDRESS: Morrison & Foerster

STREET: 755 Page Mill Road

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.10

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/720,229

FILING DATE: 26-SEP-1996

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Schiffe, J. Michael

REGISTRATION NUMBER: 40,253

REFERENCE/DOCKET NUMBER: 29938-20002.00

TELEPHONE: (415) 813-5600

TELEFAX: (415) 494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 2713 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-720-229-13

Query Match

Best Local Similarity 100.0%; Score 20; DB 3; Length 2713;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1988 TCACCACCGTCGACGACCTTC 2007

QY 1 TCACCACCGTCGACGACCTTC 20

RESULT 3

5244792-1

Patent No. 5244792

APPLICANT: BURKE, RAE L.; PACHL, CAROL; VALENZUELA, PABLO D.T.

TITLE OF INVENTION: EXPRESSION OF RECOMBINANT GLYCOPROTEIN

NUMBER OF SEQUENCES: 19

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/587,179

FILING DATE: 20-SEP-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 921,730

FILING DATE: 20-OCT-1986

APPLICATION NUMBER: 597,784

FILING DATE: 06-APR-1984

SEQ ID NO:1:

LENGTH: 3472

5244792-1

Query Match 100.0%; Score 20; DB 6; Length 3472;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCACCGTCGACGACCTTC 20

DB 2290 TCACCACCGTCGACGACCTTC 2309

RESULT 4

5244792-2

Patent No. 5244792

APPLICANT: BURKE, RAE L.; PACHL, CAROL; VALENZUELA, PABLO D.T.

TITLE OF INVENTION: EXPRESSION OF RECOMBINANT GLYCOPROTEIN

NUMBER OF SEQUENCES: 19

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/587,179

FILING DATE: 20-SEP-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 921,730

FILING DATE: 20-OCT-1986

APPLICATION NUMBER: 597,784

FILING DATE: 06-APR-1984

SEQ ID NO:2:

LENGTH: 3472

5244792-2

Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 3472;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCACCGTCGACGACCTTC 20

DB 2299 TCACCACCGTCGACGACCTTC 2318

RESULT 5

US-08-164-839-7

Sequence 7, Application US/08164839

Patent No. 5514573

GENERAL INFORMATION:

APPLICANT: YASUEDA, HISASHI

APPLICANT: NAKANISHI, KAZUO

APPLICANT: MOTOKI, MASAO

APPLICANT: NAGASE, KAZUO

APPLICANT: MATSUI, HIROSHI

TITLE OF INVENTION: GENE ENCODING TRANSGLUTAMINASE DERIVED

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESS: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,

STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/164,839

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/004,729

FILING DATE: 14-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5514573man F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 10-599-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)412-3000
TELEFAX: (703)413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2088 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: TheraGra chalcogramma
TISSUE TYPE: liver
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2085
US-08-164-839-7

Query Match 92.0%; Score 18.4; DB 1; Length 2088;
Best Local Similarity 95.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCACCACCGTCAGCACCTTC 20
Db 1802 TCACCACCGTCATCACCCTTC 1821

RESULT 6
US-08-583-799-7
Sequence 7, Application US/08583799
Patent No. 5607849
GENERAL INFORMATION:
APPLICANT: YASUEDA, HISASHI
APPLICANT: NAKANISHI, KAZUO
APPLICANT: MOTOKI, MASAO
APPLICANT: NAGASE, KAZUO
APPLICANT: MATSUI, HIROSHI
TITLE OF INVENTION: GENE ENCODING TRANSGLUTAMINASE DERIVED
TITLE OF INVENTION: FROM FISH
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/583,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/004,729
FILING DATE: 14-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5607849man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-599-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)413-2220
TELEFAX: (703)413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2088 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: TheraGra chalcogramma
TISSUE TYPE: liver
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2085
US-08-583-799-7

Query Match 92.0%; Score 18.4; DB 1; Length 2088;
Best Local Similarity 95.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCACCACCGTCAGCACCTTC 20
Db 1802 TCACCACCGTCATCACCCTTC 1821

RESULT 7
US-08-164-839-9
Sequence 9, Application US/08164839
Patent No. 5514573
GENERAL INFORMATION:
APPLICANT: YASUEDA, HISASHI
APPLICANT: NAKANISHI, KAZUO
APPLICANT: MOTOKI, MASAO
APPLICANT: NAGASE, KAZUO
APPLICANT: MATSUI, HIROSHI
TITLE OF INVENTION: GENE ENCODING TRANSGLUTAMINASE DERIVED
TITLE OF INVENTION: FROM FISH
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/164,839
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/004,729
FILING DATE: 14-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5514573man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-599-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)412-3000
TELEFAX: (703)413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2921 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: TheraGra chalcogramma
TISSUE TYPE: liver
FEATURE:

NAME/KEY: CDS
LOCATION: 32..2122
US-08-164-839-9

Query Match 92.0%; Score 18.4; DB 1; Length 2921;
Best Local Similarity 95.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCACCGTCAGCACCTTC 20
DB 1836 TCACCACCGTCATCACCCTTC 1855

RESULT 8

US-08-583-799-9
Sequence 9, Application US/08583799

Patent No. 5607849

GENERAL INFORMATION:

APPLICANT: YASUEDA, HISASHI

APPLICANT: NAKANISHI, KAZUO

APPLICANT: MOTOKI, MASAO

APPLICANT: MAGASE, KAZUO

APPLICANT: MATSUI, HIROSHI

TITLE OF INVENTION: GENE ENCODING TRANSGLUTAMINASE DERIVED

TITLE OF INVENTION: FROM FISH

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,

ADDRESSEE: P.C.

STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/583,799

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/004,729

FILING DATE: 14-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5607849man F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 10-599-0

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)412-3000

TELEFAX: (703)413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 2921 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORGANISM SOURCE:

ORGANISM: Theragra chalcogramma

TISSUE TYPE: liver

FEATURE:

NAME/KEY: CDS

LOCATION: 32..2122

US-08-583-799-9

Query Match 92.0%; Score 18.4; DB 1; Length 2921;
Best Local Similarity 95.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCACCGTCAGCACCTTC 20
DB 1836 TCACCACCGTCATCACCCTTC 1855

RESULT 9

US-09-292-034-1/c
Sequence 1, Application US/09292034

Patent No. 6492343

GENERAL INFORMATION:

APPLICANT: Reddy, P. Seethadhar

APPLICANT: Babluk, Lorne

APPLICANT: TIKOO, Suresh

TITLE OF INVENTION: PORCINE ADENOVIRUS TYPE 3 GENOME

FILE REFERENCE: 293102002400

CURRENT APPLICATION NUMBER: US/09/292,034

CURRENT FILING DATE: 1999-04-14

NUMBER OF SEQ ID NOS: 8

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1

LENGTH: 34094

TYPE: DNA

ORGANISM: Porcine Adenovirus Type 3

FEATURE:

US-09-292-034-1

Query Match 92.0%; Score 18.4; DB 4; Length 34094;
Best Local Similarity 95.0%; Pred. No. 31;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCACCGTCAGCACCTTC 20
DB 11980 TCACCACCGTCAGCACCTTC 11961

RESULT 10

US-09-142-623-10/c
Sequence 10, Application US/09142623

Patent No. 6337201

GENERAL INFORMATION:

APPLICANT: Koji YAMAI et al.

TITLE OF INVENTION: -FRUCTOFURANOSIDASE AND ITS GENE, METHOD OF

TITLE OF INVENTION: ISOLATING -FRUCTOFURANOSIDASE GENE, SYSTEM FOR PRODUCING

TITLE OF INVENTION: -FRUCTOFURANOSIDASE, AND -FRUCTOFURANOSIDASE VARIANT

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.

STREET: 2033 K Street, N.W., Suite 800

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/142,623

FILING DATE: September 10, 1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Lee Cheng

REGISTRATION NUMBER: 40,949

REFERENCE/DOCKET NUMBER: 98-0989*/LC(WMC)/144

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-721-8200

TELEFAX: 202-721-8250

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 788 base pairs

TYPE: Nucleic acid
STRANDEDNESS: Double stranded
TOPOLOGY: Linear
ORIGINAL SOURCE:
ORGANISM: Microorganism: Aspergillus niger ACE-2-1
ORGANISM: (ATCC 20611)
FEATURES:
NAME/KEY: P CDS (partial amino acid sequence)
LOCATION: 1..788
IDENTIFICATION METHOD: E
US-09-142-623-10

Query Match 84.0%; Score 16.8; DB 4; Length 788;
Best Local Similarity 90.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCACCACGCTGACGACCTTC 20
DB 772 TCACACCGCTGACACCTTC 753

RESULT 11

US-09-142-623-2/c
Sequence 2, Application US/09142623

Patent No. 6337201

GENERAL INFORMATION:

APPLICANT: KOJI YANAI et al.

TITLE OF INVENTION: -FRUCTOFRANOSIDASE AND ITS GENE, METHOD OF

TITLE OF INVENTION: ISOLATING -FRUCTOFRANOSIDASE GENE, SYSTEM FOR PRODUCING

TITLE OF INVENTION: -FRUCTOFRANOSIDASE, AND -FRUCTOFRANOSIDASE VARIANT

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.

SPRINT: 2033 K Street, N.W., Suite 800

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/142,623

FILING DATE: September 10, 1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Lee Cheng

REGISTRATION NUMBER: 40,949

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-721-8200

TELEFAX: 202-721-8250

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1905 base pairs

TYPE: Nucleic acid

STRANDEDNESS: Double stranded

TOPOLOGY: Linear

MOLECULE TYPE: Genomic DNA

ORGANISM: Microorganism: Aspergillus niger ACE-2-1

ORGANISM: (ATCC 20611)

FEATURE:

NAME/KEY: mat peptide

LOCATION: 1..1905

IDENTIFICATION METHOD: E

US-09-142-623-2

Query Match 84.0%; Score 16.8; DB 4; Length 1905;

Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TCACCACGCTGACGACCTTC 20
DB 1243 TCACACCGCTGACACCTTC 1224

RESULT 12

US-09-103-840A-2/c

Sequence 2, Application US/09103840A

Patent No. 6294328

GENERAL INFORMATION:

APPLICANT: FLEISCHMAN, Robert D.

APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

TITLE OF INVENTION: TUBERCULOSIS

FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103,840A

CURRENT FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 4403765

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis

FEATURE:

OTHER INFORMATION: CDC 1551

OTHER INFORMATION: "n" bases at various positions throughout the sequence

US-09-103-840A-2

Query Match 84.0%; Score 16.8; DB 3; Length 4403765;
Best Local Similarity 90.0%; Pred. No. 84;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCACCACGCTGACGACCTTC 20
DB 4389770 TCACCACGCTGACGACCTTC 4389751

RESULT 13

US-09-103-840A-1/c

Sequence 1, Application US/09103840A

Patent No. 6294328

GENERAL INFORMATION:

APPLICANT: FLEISCHMAN, Robert D.

APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

TITLE OF INVENTION: TUBERCULOSIS

FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103,840A

CURRENT FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 4411529

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis

OTHER INFORMATION: H37RV

US-09-103-840A-1

Query Match 84.0%; Score 16.8; DB 3; Length 4411529;
Best Local Similarity 90.0%; Pred. No. 84;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCACCACGCTGACGACCTTC 20
DB 4397533 TCACCACGCTGACGACCTTC 4397514

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2003, 19:14:25 ; Search time 72.2772 Seconds
(without alignments)
680.690 Million cell updates/sec

Title: US-09-594-065-1

Perfect score: 20
Sequence: 1 tcaccaccgtcagcacccttc 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1660708 seqs, 1229959015 residues

Total number of hits satisfying chosen parameters: 3321416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	154746	12 US-09-827-688-8	Sequence 8, Appli
2	18.4	92.0	34094	12 US-09-963-038A-1	Sequence 1, Appli
3	18.4	92.0	34094	15 US-10-199-550-1	Sequence 1, Appli
4	18.4	92.0	34094	15 US-10-245-603A-1	Sequence 1, Appli
5	16.8	84.0	84	9 US-09-864-761-31756	Sequence 31756, A
6	16.8	84.0	466	9 US-09-864-761-2089	Sequence 2089, Ap
7	16.8	84.0	498	9 US-09-864-761-15233	Sequence 15233, A
8	16.8	84.0	498	11 US-09-918-995-26919	Sequence 26919, A
9	16.8	84.0	585	13 US-10-027-632-227474	Sequence 227474,
10	16.8	84.0	788	10 US-09-990-385-10	Sequence 10, Appli
11	16.8	84.0	1336	12 US-10-017-161-569	Sequence 569, App
12	16.8	84.0	1336	12 US-10-017-161-1545	Sequence 1545, Ap
13	16.8	84.0	1339	12 US-10-017-161-813	Sequence 813, App
14	16.8	84.0	1629	14 US-10-128-714-2340	Sequence 2340, Ap
15	16.8	84.0	1691	10 US-09-938-842A-3140	Sequence 3140, Ap
16	16.8	84.0	1813	14 US-10-128-714-1340	Sequence 1340, Ap

17	16.8	84.0	1896	14 US-10-128-714-7340	Sequence 7340, Ap
18	16.8	84.0	1905	10 US-09-990-385-2	Sequence 2, Appli
19	16.8	84.0	2099	14 US-10-128-714-6340	Sequence 6340, Ap
20	16.8	84.0	3812	14 US-10-128-714-340	Sequence 340, App
21	16.8	84.0	4099	14 US-10-128-714-5340	Sequence 5340, Ap
22	16.4	82.0	628	13 US-10-027-632-289727	Sequence 289727,
23	16.4	82.0	628	13 US-10-027-632-289728	Sequence 289728,
24	16.4	82.0	628	13 US-10-027-632-289729	Sequence 289729,
25	16.4	82.0	927	9 US-09-815-242-6115	Sequence 6115, Ap
26	16.4	82.0	2394	9 US-09-815-242-4016	Sequence 4016, Ap
27	16	80.0	243	10 US-09-969-708-166	Sequence 166,
28	16	80.0	878	13 US-10-027-632-166052	Sequence 166052,
29	16	80.0	878	13 US-10-027-632-166053	Sequence 166053,
30	16	80.0	1918	10 US-09-822-863-1	Sequence 1, Appli
31	16	80.0	1918	12 US-10-281-319-1	Sequence 1, Appli
32	16	80.0	27359	10 US-09-822-863-3	Sequence 3, Appli
33	16	80.0	27359	12 US-10-281-319-3	Sequence 3, Appli
34	15.8	79.0	195	9 US-09-864-761-20885	Sequence 20885, A
35	15.8	79.0	212	11 US-09-764-891-267	Sequence 267, App
36	15.8	79.0	273	10 US-09-738-626-3003	Sequence 3003, App
37	15.8	79.0	278	9 US-09-864-761-28656	Sequence 28656, A
38	15.8	79.0	289	9 US-09-764-878-203	Sequence 203, App
39	15.8	79.0	289	14 US-10-079-854-203	Sequence 203, App
40	15.8	79.0	318	13 US-10-040-739-741	Sequence 741, App
41	15.8	79.0	323	10 US-09-878-574-2422	Sequence 2422, Ap
42	15.8	79.0	393	10 US-09-878-574-1401	Sequence 1401, Ap
43	15.8	79.0	409	9 US-09-864-761-4131	Sequence 4131, Ap
44	15.8	79.0	501	9 US-09-867-550-1533	Sequence 1533, Ap
45	15.8	79.0	549	9 US-09-864-761-12076	Sequence 12076, A

ALIGNMENTS

RESULT 1
US-09-827-688-8/c
; Sequence 8, Application US/09827688
; Publication No. US20030165476A1
; GENERAL INFORMATION:
; APPLICANT: ORSON, FRANK
; APPLICANT: KINSEY, BREMA
; APPLICANT: BHOAL, BALAIR
; TITLE OF INVENTION: MACROGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION E
; FILE REFERENCE: P01949US1/10004014
; CURRENT APPLICATION NUMBER: US/09/827,688
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/195,680
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 154746
; TYPE: DNA
; ORGANISM: HERPESVIRUS 2
US-09-827-688-8

Query Match 100.0%; Score 20; DB 12; Length 154746;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 54136 TCACCACCGTCAGCACCCTTC 54117
1 TCACCACCGTCAGCACCCTTC 20
US-09-963-038A-1/c
; Sequence 1, Application US/09963038A
; Publication No. US20030143200A1
; GENERAL INFORMATION:
; APPLICANT: Tiko, Sureeh K.
; TITLE OF INVENTION: PORCINE ADENOVIRUS E1 REGION

FILE REFERENCE: 293102003200
CURRENT APPLICATION NUMBER: US/09/963,038A
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 34094
TYPE: DNA
ORGANISM: Porcine Adenovirus Type 3
US-09-963-038A-1

Query Match 92.0%; Score 18.4; DB 12; Length 34094;
Best Local Similarity 95.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCAACGCTCAGCACCCTTC 20
DB 11980 TCACCAACGCTCAGCACCCTTC 11961

RESULT 3
US-10-199-550-1/c

Sequence 1, Application US/10199550
Publication No. US2003009615A1
GENERAL INFORMATION:
APPLICANT: TIKOO, Suresh K.
TITLE OF INVENTION: PORCINE ADENOVIRUS E1 AND E4 REGIONS
FILE REFERENCE: 293102003220
CURRENT APPLICATION NUMBER: US/10/199,550
CURRENT FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US 09/963,038
PRIOR FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 34094
TYPE: DNA
ORGANISM: Porcine Adenovirus Type 3
US-10-199-550-1

Query Match 92.0%; Score 18.4; DB 14; Length 34094;
Best Local Similarity 95.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCAACGCTCAGCACCCTTC 20
DB 11980 TCACCAACGCTCAGCACCCTTC 11961

RESULT 4
US-10-245-603A-1/c

Sequence 1, Application US/10245603A
Publication No. US2003010187A1
GENERAL INFORMATION:
APPLICANT: REDDY, Police Seshidhar
APPLICANT: TIKOO, Suresh Kumar
APPLICANT: BABIUK, Lorne A.
TITLE OF INVENTION: PORCINE ADENOVIRUS TYPE 3 GENOME
FILE REFERENCE: 293102002410
CURRENT APPLICATION NUMBER: US/10/245,603A
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: US 09/292,034
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: US 60/081,882
PRIOR FILING DATE: 1998-04-15
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 34094
TYPE: DNA
ORGANISM: Porcine Adenovirus Type 3
FEATURE:
US-10-245-603A-1

Query Match 92.0%; Score 18.4; DB 15; Length 34094;
Best Local Similarity 95.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCAACGCTCAGCACCCTTC 20
DB 11980 TCACCAACGCTCAGCACCCTTC 11961

RESULT 5

US-09-864-761-31756
Sequence 31756, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aecmics-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 31756
LENGTH: 84
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC008179.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.91
OTHER INFORMATION: EST HUMAN HIT: BB891293.1, EVALUE 4.00e-03
OTHER INFORMATION: NT HIT: L27595.1, EVALUE 1.00e-05
US-09-864-761-31756

Query Match 84.0%; Score 16.8; DB 9; Length 84;
Best Local Similarity 90.0%; Pred. No. 2.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCACACGCTGACGACCTTC 20
DB 35 TCACACGCTGACGACCTTC 54

RESULT 6

US-09-864-761-2089/c
Sequence 2089, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 2089
LENGTH: 466
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC009303.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.86
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
US-09-864-761-2089

Query Match 84.0%; Score 16.8; DB 9; Length 466;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCACACGCTGACGACCTTC 20
DB 265 TCACACGCTGACGACCTTC 246

RESULT 7

US-09-864-761-15233
Sequence 15233, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 15233
LENGTH: 498
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC008179.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.91
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.86
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
US-09-864-761-15233

Query Match 84.0%; Score 16.8; DB 9; Length 498;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DQ 1 TCACCACCGTCAAGCACCCTTC 20
|||
DB 376 TCACCACCATCAGCACCATC 395

```

RESULT 8
US-09-918-995-26919/C
Sequence 26919, Application US/09918995
Publication NO. US2003007362A1
GENERAL INFORMATION:
APPLICANT: Hysq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/05/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 26919
LENGTH: 498
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(498)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-26919

```

Query Match	84.0%	Score 16.8	DB 11	Length 498
Best Local Similarity	90.0%	Score No. 2.5e+02		
Matches 18	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY	1 TCACCACCGTTCAGCACCTTC	20		
Db	437 TCACCACCATCAGCACCTTC	418		

```

RESULT 9
US-10-027-632-227474
: Sequence 227474, Application US/10027632
: GENERAL INFORMATION:
: APPLICANT: Mang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
: POLYMORPHISMS IN THE HUMAN GENOME
: FILE REFERENCE: 108827.129
: CURRENT APPLICATION NUMBER: US/10/027,632
: CURRENT FILING DATE: 2002-04-30
: PRIOR APPLICATION NUMBER: US 60/218,006
: PRIOR FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 60/198,676
: PRIOR FILING DATE: 2000-04-20
: PRIOR APPLICATION NUMBER: US 60/193,483
: PRIOR FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: US 60/185,218
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/167,363
: PRIOR FILING DATE: 1999-11-23
: PRIOR APPLICATION NUMBER: US 60/156,358
: PRIOR FILING DATE: 1999-09-28
: PRIOR APPLICATION NUMBER: US 60/146,002
: PRIOR FILING DATE: 1999-08-09
: NUMBER OF SEQ. ID NOS: 325720
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 227474
: LENGTH: 585
: TYPE: DNA
: ORGANISM: Human
US-10-027-632-227474
: Query Match      84.0%; Score 16.8; DB 13; Length 585;

```

		Best Local Similarity	90.0%	Pred. No.	2.5e+02;	
Matches	18;	Conservative	0;	Mismatches	2;	Indels
Gaps						0;

Oy	1 TCACCACGTCAGCACCTTC	20
Db	30 TCACACTGTACGACCTTC	49

```

US-09-990-385-10/c
; Sequence 10, Application US/09990385
; Publication No. US20020192771A1
; GENERAL INFORMATION:
; APPLICANT: KOJI YANAI et al.
; TITLE OF INVENTION: BETA-FRUCTOFURANOSIDASE AND ITS GENE, METHOD OF
; ISOLATING BETA-FRUCTOFURANOSIDASE GENE, SYSTEM FOR PRODUCTION OF
; BETA-FRUCTOFURANOSIDASE, AND BETA-FRUCTOFURANOSIDASE VARIANTS
;
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/990,385
; FILING DATE: 10-Sep-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/142,623
; FILING DATE: September 10, 1998
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 2001-1611
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
;
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 788 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Double stranded
; TOPOLOGY: Linear
;
; ORIGINAL SOURCE:
; ORGANISM: Microorganism: Aspergillus niger ACE-2-1
; (ATCC 20611)
;
; FEATURE:
; NAME/KEY: P CDS (partial amino acid sequence)
; LOCATION: 1..788
; IDENTIFICATION METHOD: E
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-990-385-10
;
; Query Match 84.0%; Score 16.8; DB 10; Length 788;
; Best Local Similarity 90.0%; Pred. No. 2.5e+02;
; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; 1 TCACCACCGTCGACGACCTTC 20
; |||||||
; DB 772 TTCACCGCTCGACACTTC 753
;
; RESULT 11
; US-10-017-161-569
; Sequence 569, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:

```

```

RESULT 11
US-10-017-161-569
; Sequence 569, Application US/1001716
; Publication No. US20030143668A1
; GENERAL INFORMATION:

```

APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 084335/0152
CURRENT APPLICATION NUMBER: US/10/017,161
CURRENT FILING DATE: 2002-12-18
PRIOR APPLICATION NUMBER: JP 2001/246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2430
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 569
LENGTH: 1336
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: source
LOCATION: (1)..(1336)
FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(482)
FEATURE:
NAME/KEY: CDS
LOCATION: (515)..(645)
FEATURE:
NAME/KEY: CDS
LOCATION: (785)..(1136)
US-10-017-161-569

Query Match 84.0%; Score 16.8; DB 12; Length 1336;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCACCACCGTCAGCACCTTC 20
DB 217 TCACCACGTCAGCACCTTC 236

RESULT 12

US-10-017-161-1545
Sequence 1545, Application US/10017161
Publication No. US20030143668A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 084335/0152
CURRENT APPLICATION NUMBER: US/10/017,161
CURRENT FILING DATE: 2002-12-18
PRIOR APPLICATION NUMBER: JP 2001/246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2430
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1545
LENGTH: 1336
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: source
LOCATION: (1)..(1336)
FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(598)
FEATURE:
NAME/KEY: CDS
LOCATION: (785)..(1136)
US-10-017-161-1545

Query Match 84.0%; Score 16.8; DB 12; Length 1336;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TCACCACCGTCAGCACCTTC 20
DB 217 TCACCACGTCAGCACCTTC 236

RESULT 13

US-10-017-161-813
Sequence 813, Application US/10017161
Publication No. US20030143668A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 084335/0152
CURRENT APPLICATION NUMBER: US/10/017,161
CURRENT FILING DATE: 2002-12-18
PRIOR APPLICATION NUMBER: JP 2001/246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2430
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 813
LENGTH: 1339
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: source
LOCATION: (1)..(1339)
FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(1139)
US-10-017-161-813

Query Match 84.0%; Score 16.8; DB 12; Length 1339;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCACCACCGTCAGCACCTTC 20
DB 217 TCACCACGTCAGCACCTTC 236

RESULT 14

US-10-128-714-2340
Sequence 2340, Application US/10128714
Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Wengji
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Ercashkin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-08
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1

SEQ ID NO 2340
 LENGTH: 1629
 TYPE: DNA
 ORGANISM: Aspergillus fumigatus
 US-10-128-714-2340

Query Match 84.0%; Score 16.8; DB 14; Length 1629;
 Best Local Similarity 90.0%; Pred. No. 2.5e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCACCAACCGTCAGACCTTC 20
 DB 461 TCACCAACCGTCAGACCTTC 480

RESULT 15
 US-09-938-842A-3140/c
 Sequence 3140, Application US/0938842A
 Patent No. US20020160378A1
 GENERAL INFORMATION:
 APPLICANT: Harper, Jeff
 APPLICANT: Kieps, Joel
 APPLICANT: Wang, Xun
 APPLICANT: Zhu, Tong
 TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 TITLE OF INVENTION: SAME, AND METHODS OF USE
 FILE REFERENCE: SCRIPI300-3
 CURRENT APPLICATION NUMBER: US/09/938,842A
 CURRENT FILING DATE: 2001-08-24
 PRIOR APPLICATION NUMBER: US 60/227,866
 PRIOR FILING DATE: 2000-08-24
 PRIOR APPLICATION NUMBER: US 60/264,647
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: US 60/300,111
 PRIOR FILING DATE: 2001-06-22
 NUMBER OF SEQ ID NOS: 5379
 SEQ ID NO 3140
 LENGTH: 1691
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 US-09-938-842A-3140

Query Match 84.0%; Score 16.8; DB 10; Length 1691;
 Best Local Similarity 90.0%; Pred. No. 2.5e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCACCAACCGTCAGACCTTC 20
 DB 362 TCACCAACCGTCAGACCTTC 343

Search completed: September 16, 2003, 22:54:09
 Job time : 74.2772 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 16, 2003, 18:22:25 ; Search time 1012.28 Seconds

(without alignments)
480.194 Million cell updates/sec

Title: US-09-594-065-1

Perfect score: 20

Sequence: 1 tcaccaccgtcagcaccctc 20

Scoring table: Gapped 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estro:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_dln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	794	28	BH706301 BOMBH09TF
2	18.4	92.0	460	28	BH558118 BOMBH09TF
3	18.4	92.0	471	28	B2201556 BOMBH09TF
4	18.4	92.0	537	9	AU291667 AU291667

C 5	18.4	92.0	563	12	BM619819
C 6	18.4	92.0	574	12	BM575164
C 7	18.4	92.0	568	13	BU374492
C 8	18.4	92.0	625	12	BM635604
C 9	18.4	92.0	635	12	BM638755
C 10	18.4	92.0	658	28	BH589640
C 11	18.4	92.0	666	12	BM619759
C 12	18.4	92.0	684	12	BM586254
C 13	18.4	92.0	707	28	BM648540
C 14	18.4	92.0	784	12	B2158304
C 15	18.4	92.0	712	14	CD261546
C 16	18.4	92.0	755	13	BU374843
C 17	18.4	92.0	762	9	AL929750
C 18	18.4	92.0	774	29	B2244382
C 19	18.4	92.0	841	29	B2240124
C 20	18.4	92.0	863	29	B2260714
C 21	18.4	92.0	1106	28	B2198952
C 22	18	90.0	394	14	CA706900
C 23	17.4	87.0	138	28	BH198208
C 24	17.4	87.0	308	10	BE495478
C 25	17.4	87.0	309	9	AV549908
C 26	17.4	87.0	358	13	BU837718
C 27	17.4	87.0	368	14	CA607934
C 28	17.4	87.0	387	10	BF146112
C 29	17.4	87.0	405	13	BU822753
C 30	17.4	87.0	413	13	AV433830
C 31	17.4	87.0	413	13	BU982582
C 32	17.4	87.0	420	9	AU230462
C 33	17.4	87.0	423	14	CA712183
C 34	17.4	87.0	426	14	CA711350
C 35	17.4	87.0	429	14	CD055963
C 36	17.4	87.0	432	13	BU982723
C 37	17.4	87.0	441	10	BG417681
C 38	17.4	87.0	442	14	CD057939
C 39	17.4	87.0	452	13	BU980793
C 40	17.4	87.0	466	13	BU650160
C 41	17.4	87.0	471	28	AQ912575
C 42	17.4	87.0	474	9	AV388895
C 43	17.4	87.0	482	13	BO462866
C 44	17.4	87.0	502	10	BF145641
C 45	17.4	87.0	511	9	AU239184

ALIGNMENTS

RESULT 1
LOCUS BH706301/1
DEFINITION BOMBH09TF BO.2.3 KB Brassica oleracea genomic clone BOMBH09,
GENOMIC SURVEY SEQUENCE.
ACCESSION BH706301
VERSION BH706301.1 GI:18790010
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 794)
Tom,C.D., Van Aken,S., Uteback,T., Koo,H. and Fraser,C.M.
A whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other-GSSs: BOMBH09TF
COMMENT
Contact: Chris Town .
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1. 794
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO100DH3"
/db_xref="taxon:3712"
/clone_1ib="BO_2_3_KB"
/note="Vector: pHS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHS1 using BstXI linkers"

BASE COUNT
240 a 134 c 200 g 220 t

ORIGIN

Query Match 100.0%; Score 20; DB 28; Length 794;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCAACCTGACACCTTC 20
|||||
Db 227 TCACCAACCTGACACCTTC 208

RESULT 2
BH558118/c 460 bp DNA linear GSS 14-DEC-2001
LOCUS
DEFINITION
BOHNM39TR BOH Brassica oleracea genomic clone BOHNM39, genomic survey sequence.
ACCESSION
BH558118 GI:117809898
VERSION
KEYWORDS
SOURCE
Brassica oleracea
ORGANISM
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 460)
! eutrosids 1; Brassicales; Brassicaceae; Brassica.
REFERENCE
Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
TITLE
Whole genome shotgun sequencing of Brassica oleracea
JOURNAL
Unpublished
COMMENT
Other GSSs: BOHNM39TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1. 460
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO100DH3"
/db_xref="taxon:3712"
/clone_1ib="BOHNM39"
/note="Vector: pHS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHS1 using BstXI linkers"

BASE COUNT
117 a 92 c 126 g 125 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 28; Length 460;
Best Local Similarity 95.0%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCAACCTGACACCTTC 20
|||||
Db 122 TCACCAACCTGACACCTTC 103

RESULT 3
BZ201556/c

LOCUS
BZ201556 471 bp DNA linear GSS 11-OCT-2002
DEFINITION
CH230-399M1.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-399M1, genomic survey sequence.
ACCESSION
BZ201556
VERSION
BZ201556.1 GI:23859608
KEYWORDS
GSS.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 471)
Zhao, S., Shetty, V., Shatsman, S., Teegaye, G., Geer, K., Shvartsbeyn, A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de Jong, P. and Fraser, C.M.
TITLE
Rat BAC End Sequences from Library CHORI-230 MboI segment
JOURNAL
Unpublished
COMMENT
Other GSSs: CH230-399M1.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pje@jngmail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ordering_information.html). BAC end
page: http://www.tigr.org/cdb/bac_ends/rat/bac_end_intro.html
Plate: 399 row: M column: 1
Seq primer: T7
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1. 471
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-399M1"
/sex="female"
/cell_type="Brain"
/clone_1ib="CHORI-230 Segment 2"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"

BASE COUNT
123 a 128 c 116 g 104 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 28; Length 471;
Best Local Similarity 95.0%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCAACCTGACACCTTC 20
|||||
Db 171 TCACCAACCTGACACCTTC 152

RESULT 4
AU291667 537 bp mRNA linear EST 04-DEC-2002
LOCUS
DEFINITION
AU291667 zinnia cultured mesophyll cell equalized cDNA zinnia
elegans cDNA clone Z6332, mRNA sequence.
ACCESSION
AU291667
VERSION
AU291667.1 GI:24252175
KEYWORDS
EST.
SOURCE
Zinnia elegans
ORGANISM
Zinnia elegans
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Zinnia.

REFERENCE 1 (bases 1 to 537)
AUTHORS Demura,T., Tashiro,G., Horiguchi,G., Kishimoto,N., Kubo,M., Matsuo,K., Minami,A., Nagata-Hiwatashi,M., Nakamura,K., Okamura,Y., Sassa,N., Suzuki,S., Yazaki,J., Kikuchi,S. and Fukuda,H.
TITLE Visualization by comprehensive microarray analysis of gene expression programs during transdifferentiation of mesophyll cells into xylem cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (24), 15794-15799 (2002)
COMMENT Contact: Taku Demura
Morphogenesis Research Group
RIKEN Plant Science Center
1-7-22 Suehirocho, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9605
Fax: 81-45-503-9573
Email: demura@postman.riken.go.jp
This clone was obtained at our laboratory.
Seq primer: M13 reverse.

FEATURES
SOURCE
1..537
/organism="Zinnia elegans"
/mol_type="mRNA"
/cultivar="Canary bird"
/db_xref="taxon:34245"
/clone="Z6332"
/tissue_type="mesophyll cell"
/clone_1ib="zinnia cultured mesophyll cell equalized cDNA"
/note="Vector: pGEM-T easy; cultured in tracheary element differentiation-inductive medium"

BASE COUNT 177 a 158 c 90 g 105 t 7 others
ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 537;
Best Local Similarity 95.0%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCACCGTCAGCACCCTTC 20
|||||
Db 312 TCACCACCGTCAGCACCCTTC 331

RESULT 5
BM619819/c 563 bp mRNA linear EST 25-FEB-2002
LOCUS 17000687440937 A.Gam.ad.cDNA1 Anopheles gambiae cDNA clone
DEFINITION 19600449618554 5', mRNA sequence.
ACCESSION BM619819
VERSION BM619819.1 GI:18918237
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anophelinae.
REFERENCE 1 (bases 1 to 563)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE Celera Anopheles gambiae EST project
JOURNAL Unpublished
COMMENT Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltR@celera.com
Plate: NU01004N8 row: H column: 08
Seq primer: M13 Reverse.

FEATURES
SOURCE
1..563
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std. chromosome)"
/db_xref="taxon:7165"
/clone_1ib="A.Gam.ad.cDNA.blood1"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Whole adult mosquitoes (mixed sex) frozen on liquid nitrogen 24 hours after human blood feeding. cDNA inserts >500 bp cloned directionally into pSPORT 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)"

BASE COUNT 175 a 139 c 136 g 118 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 568;
Best Local Similarity 95.0%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCACCGTCAGCACCCTTC 20

RESULT 6
BM575164/c 568 bp mRNA linear EST 22-FEB-2002
LOCUS 17000659210616 A.Gam.ad.cDNA.blood1 Anopheles gambiae cDNA clone
DEFINITION 19600447014826 5', mRNA sequence.
ACCESSION BM575164
VERSION BM575164.1 GI:18863631
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anophelinae.
REFERENCE 1 (bases 1 to 568)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE Celera Anopheles gambiae EST project
JOURNAL Unpublished
COMMENT Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltR@celera.com
Plate: NU01003S1T row: G column: 09
Seq primer: M13 Reverse.

FEATURES
SOURCE
1..568
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std. chromosome)"
/db_xref="taxon:7165"
/clone_1ib="A.Gam.ad.cDNA.blood1"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Whole adult mosquitoes (mixed sex) frozen on liquid nitrogen 24 hours after human blood feeding. cDNA inserts >500 bp cloned directionally into pSPORT 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)"

BASE COUNT 175 a 139 c 136 g 118 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 568;
Best Local Similarity 95.0%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCACCGTCAGCACCCTTC 20

Db 61 TCACCTCGTCAGCACCCTTC 42

|||||

RESULT 7
BU374492 574 bp mRNA linear EST 28-NOV-2002
LOCUS 603813877F1 CSEQCHN74 Gallus gallus cDNA clone CHEST804j20 5', mRNA
DEFINITION sequence.

ACCESSION BU374492
VERSION BU374492.1 GI:25862493
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archaeopteryx; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 574)
Boardman, P. B., Sanz-Ezquerro, J., Overton, I. M., Burt, D. W., Boesch, E.,
Fong, M. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
A Comprehensive Collection of Chicken CDNA
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
MEDLINE 12445392
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

FEATURES
source
1..574
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST804j20"
/sex="Female"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="CSEQCHN74"
/note="Organ: kidney + adrenal; Vector: pBluescript II
KS(+); Site 1: EcoRI; Site 2: NotI; This normalized
library was constructed from 1 million independent clones.
cDNA synthesis was initiated using an oligo(dT) primer,
using methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with EcoRI
size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., EMAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

BASE COUNT 152 a 140 c 170 g 112 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 13; Length 574;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCAACGTCAGCACCCTTC 20
|||||

Db 211 TCACCAACGTCAGCACCCTTC 230
|||||

RESULT 8
BM635604/c 625 bp mRNA linear EST 26-FEB-2002
LOCUS 17000687558989 A.Gam.ad.cDNA1 Anopheles gambiae cDNA clone
DEFINITION

19600449658640 5', mRNA sequence.

ACCESSION BM635604
VERSION BM635604.1 GI:18935115
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
1 (bases 1 to 625)
Holt, R. A., Lin, J. -J., Murphy, S. D., Evans, C. A., Kraft, C. L., Charlab
R., Collins, F. H., Venter, J. C. and Hoffman, S. L.
Celera Anopheles gambiae EST project
Unpublished
Contact: Holt R. A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltR@celera.com
Place: NU01004HNJ row: N column: 14
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES
source
1..625
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-5T (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449658640"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="A.Gam.ad.cDNA1"
/note="Vector: pSport1; Site 1: SalI; Site 2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen.
cDNA inserts >500 bp cloned directionally into pSport 1.
Not 1 site is 3'. Clones available through the Malaria
Research and Reference Reagent Resource Center
(www.malaria.mr4.org)."

BASE COUNT 140 a 167 c 180 g 138 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 625;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCAACGTCAGCACCCTTC 20
|||||

Db 307 TCACCAACGTCAGCACCCTTC 288
|||||

RESULT 9
BM638755/c 635 bp mRNA linear EST 26-FEB-2002
LOCUS 17000687278313 A.Gam.ad.cDNA1 Anopheles gambiae cDNA clone
DEFINITION 1960044961587 5', mRNA sequence.

ACCESSION BM638755
VERSION BM638755.1 GI:18938278
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
1 (bases 1 to 635)
Holt, R. A., Lin, J. -J., Murphy, S. D., Evans, C. A., Kraft, C. L., Charlab
R., Collins, F. H., Venter, J. C. and Hoffman, S. L.
Celera Anopheles gambiae EST project
Unpublished
Contact: Holt R. A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151

FEATURES
source
Fax: 2404534580
Email: HoltR@celera.com
Plate: NU01004N8P row: P column: 01
Seq primer: M13 Reverse.

Location/Qualifiers
1. 635
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std. chromosome)"
/db_xref="taxon:7165"
/clone="19600449617587"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_lib="A. Gam. ad. cDNA1"
/note="Vector: pSport1; Site 1: SalI; Site 2: NotI; whole adult mosquitoes (mixed sex) frozen on liquid nitrogen. cDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)."

BASE COUNT
ORIGIN
164 a 145 c 221 g 105 t

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 12; Length 635;
Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY
1 TCACCACTGTCAGCACCTTC 20
|||||
Db 477 TCACCTCGTCAGCACCTTC 458

RESULT 10
LOCUS
BH689640 658 bp DNA linear GSS 19-FEB-2002
DEFINITION
BOMGX07R BO_2_3 KB Brassica oleracea genomic clone BOMGX07,
genomic survey sequence.
ACCESSION
BH689640
VERSION
BH689640.1 GI:18760077
KEYWORDS
GSS.
SOURCE
Brassica oleracea
ORGANISM
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid1 II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 658)
Town,C.D., Van Aken,S., Uteerback,T., Koo,H. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other_GSSs: BOMGX07TF
COMMENT
Contact: Chris Town
TIOR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: ctown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1. 658
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO100DH3"
/db_xref="taxon:3712"
/clone="BOMGX07"
/clone_lib="BO_2_3 KB"
/note="Vector: pHO51; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHO51 using BstXI linkers"

BASE COUNT
ORIGIN
154 a 211 c 142 g 151 t

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 28; Length 658;
Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY
1 TCACCACTGTCAGCACCTTC 20
|||||
Db 309 TCACCACTGTCAGCACCTTC 328

RESULT 11
LOCUS
BM619759/c 666 bp mRNA linear EST 25-FEB-2002
DEFINITION
170006873440855 A.Gam.ad.cDNA1 Anopheles gambiae cDNA clone
19600449618598 5', mRNA sequence.
BM619759
BM619759.1 GI:18918177

ACCESSION
BM619759
VERSION
BM619759.1
KEYWORDS
EST.
SOURCE
Anopheles gambiae (African malaria mosquito)
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
1 (bases 1 to 666)
Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
Celera Anopheles gambiae EST project
Unpublished
Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltR@celera.com
Plate: NU01004N8P row: J column: 04
Seq primer: M13 Reverse.

FEATURES
source
Location/Qualifiers
1. 666
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std. chromosome)"
/db_xref="taxon:7165"
/clone="19600449618598"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_lib="A. Gam. ad. cDNA1"
/note="Vector: pSport1; Site 1: SalI; Site 2: NotI; whole adult mosquitoes (mixed sex) frozen on liquid nitrogen. cDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)."

BASE COUNT
ORIGIN
188 a 156 c 192 g 130 t

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 12; Length 666;
Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY
1 TCACCACTGTCAGCACCTTC 20
|||||
Db 229 TCACCTCGTCAGCACCTTC 210

RESULT 12
LOCUS
BM586254 684 bp mRNA linear EST 25-FEB-2002
DEFINITION
17000687312069 A.Gam.ad.cDNA1 Anopheles gambiae cDNA clone
19600449707650 5', mRNA sequence.
BM586254
BM586254.1 GI:18882115

ACCESSION
BM586254
VERSION
BM586254.1
KEYWORDS
EST.
SOURCE
Anopheles gambiae (African malaria mosquito)

ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.

REFERENCE 1 (bases 1 to 684)
AUTHORS Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab, R., Collins, F.H., Venter, J.C. and Hoffman, S.L.
TITLE Celera Anopheles gambiae EST project
JOURNAL Unpublished
COMMENT Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU010049VI row: H column: 04
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES
source 1..684
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std. chromosome)"
/db_xref="taxon:7165"
/clone="19600449707650"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_1lb="A.Gam.ad.cDNA.bloodi"
/note="Vector: pSport1; Site 1: SalI; Site 2: NotI; Whole adult mosquitoes (mixed sex) frozen on liquid nitrogen 24 hours after human blood feeding. cDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)"

BASE COUNT 182 a 158 c 232 g 112 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 684;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCAACGCTCAGCACCTTC 20
|||||
523 TCACCTCCGCTCAGCACCTTC 504

RESULT 13
BM648540 687 bp mRNA linear EST 26-FEB-2002
LOCUS 17000687326166 A.Gam.ad.cDNA1 Anopheles gambiae CDNA clone
DEFINITION 19600449658826 5', mRNA sequence.
ACCESSION BM648540
VERSION BM648540.1 GI:18948051
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.

REFERENCE 1 (bases 1 to 687)
AUTHORS Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab, R., Collins, F.H., Venter, J.C. and Hoffman, S.L.
TITLE Celera Anopheles gambiae EST project
JOURNAL Unpublished
COMMENT Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU010049VI row: F column: 08
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES

source 1..687
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std. chromosome)"
/db_xref="taxon:7165"
/clone="19600449658826"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_1lb="A.Gam.ad.cDNA1"
/note="Vector: pSport1; Site 1: SalI; Site 2: NotI; Whole adult mosquitoes (mixed sex) frozen on liquid nitrogen 24 hours after human blood feeding. cDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)."

BASE COUNT 175 a 157 c 238 g 117 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 687;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACCAACGCTCAGCACCTTC 20
|||||
546 TCACCTCCGCTCAGCACCTTC 527

RESULT 14
B2158304 704 bp DNA linear GSS 11-OCT-2002
LOCUS CH230-38714.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
DEFINITION CH230-38714, genomic survey sequence.
ACCESSION B2158304
VERSION B2158304.1 GI:23799259
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 704)
AUTHORS Zhao, S., Shetty, D., Shatsman, S., Tsagaye, G., Geer, K., Shvartsbeyn, A., Gebregiorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de Jong, P. and Fraser, C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 MboI segment
JOURNAL Unpublished
COMMENT Other_GSSs: CH230-38714.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230 (http://www.chori.org/bacpac/rat230.htm). For BAC library availability, please contact pletier at Jong (pletier@jgmi.choi.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or eting_information.htm). BAC end page: http://www.tigr.org/cdb/bac_ends/rat/bac_end_intro.html
Plate: 387 row: I column: 4
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers

FEATURES
source 1..704
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-38714"
/sex="Female"
/cell_type="Brain"

```

/clone.lib="CHORI-230 Segment 2"
/notes="vector: pTABACT1.3; Site_1: MboI; Site_2: MboI;
CHORI-230 Rat (BN/SNNHsd/MCW) BAC library produced by
Pieter de Jong"

```

```

BASE COUNT      156 a      171 c      204 g      173 t
ORIGIN

```

```

Query Match      92.0%; Score 18.4; DB 28; Length 704;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 TCACCAACGTCGACACCTTC 20
        |||||
Db      74 TCACCAACGTCGACACCTTC 93

```

RESULT 15

```

LOCUS      CD261546                      712 bp      mRNA      linear      EST 23-MAY-2003
DEFINITION pSMA012x015f_189662 sMA: Phytophthora sojae grown in synthetic
ACCESSION  CD261546
VERSION     CD261546
KEYWORDS    CD261546.1 GI:31049373
SOURCE      EST.
ORGANISM    Phytophthora sojae
            Phytophthora sojae
            Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
            Phytophthora.

```

```

REFERENCE   1 (bases 1 to 712)
AUTHORS     Tyler,B.M., Judelson,H.S., Gijzen,M., Dean,R.A. and Waugh,M.E.
TITLE       USDA-IRAS: Expression of Phytophthora sojae genes during infection
            and propagation
JOURNAL     Unpublished
COMMENT     Contact: Tyler B
            VBI
            1880 Pratt Dr., Blacksburg, VA 24061, USA
            Tel: 540-231-7318
            Email: bmt Tyler@vt.edu

```

```

PCR Primers
FORWARD: BK reverse
Plate: 012 row: 0 column: 15
Seq primer: BK reverse
High quality sequence stop: 712.
Location/Qualifiers
1..712

```

```

FEATURES
source

```

```

/organism="Phytophthora sojae"
/mol_type="mRNA"
/db_xref="taxon:67593"
/clone="sMA012015"
/tissue_type="mycelium"
/cell_line="P6497"
/dev_stage="mycelium"
/clone.lib="sMA: Phytophthora sojae grown in synthetic
medium"
/notes="vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

```

```

BASE COUNT      172 a      203 c      213 g      124 t
ORIGIN

```

```

Query Match      92.0%; Score 18.4; DB 14; Length 712;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 TCACCAACGTCGACACCTTC 20
        |||||
Db      693 TCACCAACGTCGACACCTTC 712

```

```

Search completed: September 16, 2003, 20:39:49
Job time : 1016.28 secs

```

BEST AVAILABLE COPY